99/100

Science Policy Study—Hearings Volume 5
THE FUTURE OF SCIENCE

NEWARK PUBLIC LIBRARY

SEP 12 1986

U. S. GOVT. PUBS.
DEPOSITORY 376

HEARING
BEFORE THE
TASK FORCE ON SCIENCE POLICY
OF THE
COMMITTEE ON
SCIENCE AND TECHNOLOGY
HOUSE OF REPRESENTATIVES
NINETY-NINTH CONGRESS
FIRST SESSION

MAY 2, 1985

[No. 100]

Printed for the use of the
Committee on Science and Technology

Y4.Sci 2:

READEX
FICHE
best copy available

Science Policy Study—Hearings Volume 5
THE FUTURE OF SCIENCE

HEARING

BEFORE THE

TASK FORCE ON SCIENCE POLICY

OF THE

COMMITTEE ON
SCIENCE AND TECHNOLOGY
HOUSE OF REPRESENTATIVES

NINETY-NINTH CONGRESS

FIRST SESSION

MAY 2, 1985

[No. 100]

Printed for the use of the
Committee on Science and Technology

U.S. GOVERNMENT PRINTING OFFICE
52-240 O
WASHINGTON : 1986

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, DC 20402

52-240 O - 86 - 1

COMMITTEE ON SCIENCE AND TECHNOLOGY

DON FUQUA, Florida, *Chairman*

ROBERT A. ROE, New Jersey
GEORGE E. BROWN, Jr., California
JAMES H. SCHEUER, New York
MARILYN LLOYD, Tennessee
TIMOTHY E. WIRTH, Colorado
DOUG WALGREN, Pennsylvania
DAN GLICKMAN, Kansas
ROBERT A. YOUNG, Missouri
HAROLD L. VOLKMER, Missouri
BILL NELSON, Florida
STAN LUNDINE, New York
RALPH M. HALL, Texas
DAVE McCURDY, Oklahoma
NORMAN Y. MINETA, California
MICHAEL A. ANDREWS, Texas
BUDDY MacKAY, Florida**
TIM VALENTINE, North Carolina
HARRY M. REID, Nevada
ROBERT G. TORRICELLI, New Jersey
RICK BOUCHER, Virginia
TERRY BRUCE, Illinois
RICHARD H. STALLINGS, Idaho
BART GORDON, Tennessee
JAMES A. TRAFICANT, Jr., Ohio

MANUEL LUJAN, Jr., New Mexico*
ROBERT S. WALKER, Pennsylvania
F. JAMES SENSENBRENNER, Jr., Wisconsin
CLAUDINE SCHNEIDER, Rhode Island
SHERWOOD L. BOEHLERT, New York
TOM LEWIS, Florida
DON RITTER, Pennsylvania
SID W. MORRISON, Washington
RON PACKARD, California
JAN MEYERS, Kansas
ROBERT C. SMITH, New Hampshire
PAUL B. HENRY, Michigan
HARRIS W. FAWELL, Illinois
WILLIAM W. COBEY, Jr., North Carolina
JOE BARTON, Texas
D. FRENCH SLAUGHTER, Jr., Virginia
DAVID S. MONSON, Utah

Harold P. Hanson, *Executive Director*
Robert C. Ketcham, *General Counsel*
Regina A. Davis, *Chief Clerk*
Joyce Gross Freiwald, *Republican Staff Director*

Science Policy Task Force

DON FUQUA, Florida, *Chairman*

GEORGE E. BROWN, Jr., California
TIMOTHY E. WIRTH, Colorado
DOUG WALGREN, Pennsylvania
HAROLD L. VOLKMER, Missouri
STAN LUNDINE, New York
NORMAN Y. MINETA, California
HARRY M. REID, Nevada
RICK BOUCHER, Virginia
RICHARD H. STALLINGS, Idaho

MANUEL LUJAN, Jr., New Mexico*
ROBERT S. WALKER, Pennsylvania
F. JAMES SENSENBRENNER, Jr., Wisconsin
CLAUDINE SCHNEIDER, Rhode Island
SHERWOOD L. BOEHLERT, New York
TOM LEWIS, Florida
SID W. MORRISON, Washington
RON PACKARD, California

John D. Holmfeld, *Study Director*
R. Thomas Weimer, *Republican Staff Member*

---

*Ranking Republican Member.
**Serving on Committee on the Budget for 99th Congress.

CONTENTS

WITNESSES

Page
May 2, 1985:
  Dr. Frank Press, president, National Academy of Sciences, Washington, DC 3
  Prepared statement 5
  Dr. Leon T. Silver, William A. Keck Foundation Professor of Geology, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA; and chairman, Committee on Science, Engineering, and Public Policy, National Academy of Sciences 8, 42
  Prepared statement 46
  Dr. Jacob Bigeleisen, leading professor, Department of Chemistry, State University of New York at Stony Brook, Stony Brook, NY; and member, Committee on Science, Engineering, and Public Policy, National Academy of Sciences 9
  Prepared statement 13
  Dr. Floyd E. Bloom, director and member, Division of Preclinical Neuroscience and Endocrinology, Scripps Clinic and Research Foundation, LaJolla, CA; and member, Committee on Science, Engineering, and Public Policy, National Academy of Sciences 20
  Prepared statement 23
  Dr. Edward A. Mason, vice president, research, Amoco Corp., Naperville, IL; and member, Committee on Science, Engineering, and Public Policy, National Academy of Sciences 29
  Prepared statement 31
  Dr. Herbert A. Simon, Richard King Mellon University Professor of Computer Science and Psychology, Carnegie-Mellon University, Pittsburgh, PA; and member, Committee on Science, Engineering, and Public Policy, National Academy of Sciences 36
  Prepared statement 38
  Hon. Emilio Q. Daddario, Esq., Wilkes, Artis, Hedrick & Lane, attorneys at law, Washington, DC; former Member of Congress from the State of Connecticut; and member, Committee on Science, Engineering, and Public Policy, National Academy of Sciences 55
  Prepared statement 56
  Panel discussion 57
Appendix 1: Questions and answers for the record:
  Dr. Jacob Bigeleisen 77
  Dr. Floyd E. Bloom 83
  Dr. Herbert A. Simon 87
Appendix 2: Additional materials for the record:
  The Outlook for Science and Technology 1985 95
    Preface 97
    Overview 105
    Recent progress in science and technology 109
      Oncogenes 109
      Atherosclerosis 110
      Parasitism 112
      Chemical and process engineering for biotechnology 113
      Advance polymeric composites 114
      Supercomputer architectures 115
      Information technology in precollege eduction 116
      Opportunities in physics 117

Appendix 2—Continued

The Outlook for Science and Technology 1985—Continued

An outline of selected issues—Continued

Page

- International competition in science and technology 121
  - Supercomputers 121
  - Biochemical engineering 123
  - Advanced polymeric composites 125
  - Issues for the Congress 125
- Scientific and engineering personnel 126
  - Starting a research career 126
  - Clinicians in research 127
  - Possible shortages 127
  - Doctorates for non-U.S. citizens 128
  - Issues for the Congress 130
- Cooperative work across disciplines 131
  - The special case of agriculture 132
  - Issues for the Congress 133
- Research and Transportation 133
  - Highways 134
  - Issues for the Congress 135
- Facilities and instrumentation 136
  - Large-scale facilities 136
  - Medium-scale instrumentation 137
  - Planning new facilities 140
  - Issues for the Congress 141
- Issues in genetic engineering 143
  - Issues for the Congress 144
- Issues in human biology 144
  - In vitro fertilization and implantation of human eggs 145
  - Genetic therapy of human cells 145
  - Issues for the Congress 147
- Scientific communication, technology transfer, and national security 147
  - The Department of Defense 148
  - The Department of Commerce 148
  - Issues for the Congress 150
- Global atmospheric effects of nuclear war 151
  - Issues for the Congress 152
- Final comment 152

Research Briefings for the Office of Science and Technology Policy, the National Science Foundation, and Selected Federal Departments and Agencies, Committee on Science, Engineering, and Public Policy, National Academy of Sciences, National Academy of Engineering, Institute of Medicine, National Academy Press, Washington, DC, 1983 154

- Preface 158
- Report of the Research Briefing Panel on Mathematics 161
- Report of the Research Briefing Panel on Atmospheric Sciences 178
- Report of the Research Briefing Panel on Astronomy and Astrophysics 189
- Report of the Research Briefing Panel on Agricultural Research 213
- Report of the Research Briefing Panel on Neuroscience 234
- Report of the Research Briefing Panel on Human Health Effects of Hazardous Chemical Exposures 250
- Report of the Research Briefing Panel on Materials Science 261

Research Briefings 1983 for the Office of Science and Technology Policy, the National Science Foundation, and Selected Federal Departments and Selected Agencies, Committee on Science, Engineering, and Public Policy, National Academy of Sciences, National Academy of Engineering, Institute of Medicine, National Academy Press, Washington, DC, 1983 282

- Preface 285
- Report of the Research Briefing Panel on the Solid Earth Sciences 288
- Report of the Research Briefing Panel on Cognitive Science and Artificial Intelligence 305
- Report of the Research Briefing Panel on Immunology 323
- Report of the Research Briefing Panel on Computers in Design and Manufacturing 334

Appendix 2—Continued

Research Briefings 1983—Continued

Page

Report of the Research Briefing Panel on Computers in Design and Manufacturing 334

Report of the Research Briefing Panel on Selected Opportunities in Chemistry 349

Research Briefings 1984 for the Office of Science and Technology Policy, the National Science Foundation, and Selected Federal Departments and Agencies, Committee on Science, Engineering, and Public Policy, National Academy of Sciences, National Academy of Engineering, Institute of Medicine, National Academy Press, Washington, DC 1984 375

Preface 379

Report of the Research Briefing Panel on Computer Architecture 382

Report of the Research Briefing Panel on Information Technology in Precollege Education 396

Report of the Research Briefing Panel on Chemical and Process Engineering for Biotechnology 410

Report of the Research Briefing Panel on High-Performance Polymer Composites 423

Report of the Research Briefing Panel on the Biology of Oncogenes 435

Report of the Research Briefing Panel on Interactions Between Blood and Blood Vessels (Including the Biology of Atherosclerosis) 446

Report of the Research Briefing Panel on the Biology of Parasitism 457

Report of the Research Briefing Panel on Solar-Terrestrial Plasma Physics 467

Report of the Research Briefing Panel on Selected Opportunities in Physics 480

Research Briefings 1985 for the Office of Science and Technology Policy, the National Science Foundation, and Selected Federal Departments and Agencies, Committee on Science, Engineering, and Public Policy, National Academy of Sciences, National Academy of Engineering, Institute of Medicine, National Academy Press, Washington, DC, 1985 492

Preface 496

Report of the Research Briefing Panel on Remote Sensing of the Earth 500

Report of the Research Briefing Panel on Pain and Pain Management 517

Report of the Research Briefing Panel on Biotechnology in Agriculture 531

Report of the Research Briefing Panel on Weather Prediction Technologies 544

Report of the Research Briefing Panel on Ceramics and Ceramic Composites 556

Report of the Research Briefing Panel on Scientific Frontiers and the Superconducting Super Collider 569

Report of the Research Briefing Panel on Computer Vision and Pattern Recognition 582

Appendix 2—Continued

[illegible]

Report of the [illegible]

Report of the [illegible]

Report of the [illegible]

THE FUTURE OF SCIENCE

WEDNESDAY, MAY 2, 1985

House of Representatives,
Committee on Science and Technology,
Task Force on Science Policy,
Washington, DC.

The task force met, pursuant to notice, at 9:36 a.m., in room 2318, Rayburn House Office Building, Hon. Don Fuqua (chairman of the task force) presiding.

Mr. Fuqua. Today the Science Policy Task Force takes up a subject which is central to the future of American national science policy, namely the future of American science.

Our agenda includes a wide range of policy issues, such as funding levels, funding mechanisms, institutional roles, and educational practices. We must not lose sight, however, of the fact that the primary purpose of all those policies, practices, and procedures is the conduct of scientific research.

An important question, therefore, is what are the future longer term trends within science itself? What do these trends mean for our science policies? What will remain constant and what is likely to change in the content and practice of science, and how must our policies be adjusted to take those developments into account?

This assessment of the future of American science is being done within an environment which is surely changing. Two of the major factors in that change are the significantly different economic situation of the United States and the emergence, in several other countries, of major areas of strength in science.

The new economic situation in this country includes the severe constraints on all areas of the Federal budget and a switch from high to low rates of inflation. The emergence of strong science communities abroad raises important questions about the proper balance between competition and cooperation in international science. Both developments raise difficult questions of priorities among the many opportunities in science before us.

To discuss the future developments within science itself, we are fortunate to have with us today a group of distinguished scientists from the Committee on Science, Engineering and Public Policy of the National Academy of Sciences.

I take this occasion to note that the last time the Academy's Public Policy Committee appeared to testify in this room, its chairman was Dr. George Low. Throughout his long and distinguished career, Dr. Low made many significant contributions to America's space program, to American higher education, and to American science policy.

The untimely loss of George was a loss to all of us, and we must count ourselves fortunate that his work in the National Academy is so ably carried on by his former colleagues under the leadership of Dr. Leon Silver.

The Public Policy Committee has, over the last 6 or 7 years, been engaged in a close examination of the future prospects of American science. The vehicle for these examinations has been the biannual *Five-Year Outlook for Science and Technology* and the annual research briefings for the Office of Science and Technology Policy.

We have asked Dr. Silver and his committee to distill for us, insofar as that is possible, the conclusions, opportunities and trends which they foresee based on those reports.

To lead off the Academy's presentation, we welcome back to the committee Dr. Frank Press, the president of the National Academy of Sciences. We have had many fruitful contacts with Dr. Press in the various high offices that he has held, including membership on the National Science Board, Science Advisor to President Carter, and now, president of the National Academy of Sciences.

We wish to welcome all of you.

Before calling on Dr. Press, I would like to call on our ranking minority member of the task force and the full committee, Congressman Manuel Lujan.

Mr. LUJAN. Thank you, Mr. Chairman.

I don't have an opening statement. I am just pleased to be here this morning and go over the report of the group. I look forward to listening to all of the comments by all of the members, and perhaps we will have some questions later on.

Thank you very much, Mr. Chairman.

Mr. FUQUA. Thank you.

Frank, we will be delighted to hear from you now.

[A biographical sketch of Dr. Press follows:]

DR. FRANK PRESS

Frank Press was born in Brooklyn, New York in 1924. He received his undergraduate degree in physics from the City College of New York, and advanced degrees in geophysics from Columbia University in 1946 and 1949, when he joined the Columbia faculty, becoming associate professor in 1952, working in the areas of geophysics and oceanography. In 1955 Dr. Press was appointed professor of geophysics at the California Institute of Technology, and two years later became director of its Seismological Laboratory. He was named in 1965 as the head of the then Department of Geology and Geophysics at the Massachusetts Institute of Technology [MIT], which under his leadership, expanded into planetary sciences, oceanography, interdisciplinary studies, and the joint program with the Woods Hole Oceanographic Institution, and was renamed the Department of Earth and Planetary Sciences. In 1977 he was appointed by President Carter as the President's Science Advisor and Director of the Office of Science and Technology Policy. In January, 1981, he returned to MIT where he was appointed Institute Professor, a title MIT reserves for scholars of special distinction. Dr. Press returned to Washington in July 1981 as the 19th President of the National Academy of Sciences, elected by its members to a six-year term.

Dr. Press is recognized internationally for his pioneering contributions in geophysics, oceanography, lunar and planetary sciences, and natural resource exploration, but his primary scientific activities have been in seismology and the study of the earth's deep interior. Recognizing the importance of long-period surface waves in studying the earth's structure, he developed the theory for these waves and the instrumentation to record them. Today the analyses of seismic surface waves and free oscillations are among the most powerful techniques for studying the structure and properties of the earth's crust and deep interior. Dr. Press also saw the need to develop techniques for geophysical studies of the moon and planets, using landed observatories. Author of 160 scientific papers, he is also the co-author of the textbook *Earth*, widely used in courses in both American and foreign universities.

Dr. Press has been a leader in major national and international projects. He helped organize and gave impetus to the International Geophysical Year, the first coordinated worldwide attempt to measure and map various geophysical phenomena, a decade-long effort that involved international explorations of Antarctica and the oceans. Mt. Press in Antarctica is named for him. Dr. Press provided leadership in research efforts on earthquake prediction in the United States, and in international cooperation with Japan, the USSR, and the People's Republic of China.

As NAS President, Dr. Press will continue a long career of public service, in additon to his distinguished scientific work. He served on the President's Science Advisory Committee during the Kennedy Administration and on the Baker and Ramo Presidential Advisory Committee during the Ford Administration. He was appointed by President Nixon to the National Science Board, which is the policy-making body of the National Science Foundation, and he also served on the Lunar and Planetary Missions Board of the National Aeronautics and Space Administration. Dr. Press participated in the bilateral science agreement negotiations with the Soviet Union, and was a member of the U.S. delegation to nuclear test ban negotiations in Geneva and Moscow.

Major initiatives of his Washington service as OSTP Director and Science Advisor during the Carter Administration included increasing the Federal commitment to the support of basic research; the introduction of new measures to spur industrial innovation; joint research ventures involving industry, the university, and the Government; and regulatory reform, particularly in improving the scientific basis of proposed regulations. Dr. Press was largely responsible for the U.S.-China scientific cooperation agreements in 1979.

Dr. Press is a member of several professional organizations, and is a former President of both the Seismological Society of America and the American Geophysical Union. He was elected to the National Academy of Sciences in 1958, the American Academy of Arts and Sciences in 1966, and the American Philosophical Society. In 1981 he was elected as a foreign member of the French Academy of Sciences, and to the Board of Trustees of both the Sloan Foundation and Rockefeller University, as well as to the membership of the Corporation of the Massachusetts Institute of Technology. He is the recipient of numerous honors, among which are the Gold Medal of the Royal Astronomical Society, the Arthur L. Day Medal of the Geological Society, the Bowie Medal of the American Geophysical Union, and in 1982, the Maurice Ewing Medal of the Society of Exploratory Geophysicists. He was awarded the Department of the Interior's Public Service Award in 1971 and NASA's Distinguished Public Service Medal in 1973. In 1982, during the 9th Annual Survey of American Leaders conducted by *U.S. News and World Report*, and in 1984, during the 11th Annual Survey, he was named the most influential man in science. Dr. Press has received 17 honorary doctoral degrees. His unique distinction lies perhaps in the dual contribution of the impact of his scientific work on the development of moden geophysicis and the influence of his personal leadership in national science planning and administration.

Dr. Press is married to the former Billie Kallick of St. Louis. The Presses have two children and two grandchildren.

STATEMENT OF DR. FRANK PRESS, PRESIDENT, NATIONAL ACADEMY OF SCIENCES, WASHINGTON, DC

Dr. Press. Thank you, Mr. Chairman.

I will be very brief this morning in order for our committee here to begin its briefing.

The occasion of this briefing is the report just off the press this week, *The Outlook for Science and Technology, 1985.* As you know, this report flows directly from this very committee of Congress, going back some years.

The report is prepared by our Committee on Science, Engineering, and Public Policy, which perhaps is the most important committee in the National Academy complex. It is a joint committee of the Academies of Science and Engineering, and the Institute of Medicine.

In preparing this report, we benefited greatly from a discussion with you, Senator Gorton, other members of this committee and its staff, in reorganizing our process to highlight and point up those issues which result from scientific developments which should be on the congressional agenda.

We have taken pains in this report to highlight particularly for your attention problem areas, opportunities, but particularly those items that you may wish to consider, as the congressional committee responsible for these very important areas in science and technology. So the policy issues have been pointed up in this particular report, more so than ever before.

[The prepared statement of Dr. Press follows:]

PREPARED STATEMENT OF FRANK PRESS, PRESIDENT, NATIONAL ACADEMY OF SCIENCES

Good morning. As you can see, in contrast to my previous appearances before this committee, I have lots of company this time. Therefore, I'll keep my introductory remarks quite brief, but do hope to involve myself in the follow-up discussion. The subject this morning -- that of the future of science -- is a broad one. However we have a specific focus for that topic: the 1985 Outlook for Science and Technology.

The Outlook was prepared under the aegis of COSEPUP -- the Committee on Science, Engineering, and Public Policy. COSEPUP is a joint committee of the two Academies of sciences and engineering and of the Institute of Medicine. All of the gentlemen joining me this morning are members of COSEPUP. And, as such, all were involved in preparing the 1985 Outlook, and several in the previous Outlooks. They have also been involved in preparing the research briefings on selected fields of science and technology, done under COSEPUP guidance for OSTP, NSF, and other federal agencies.

As such, the witnesses this morning can speak not only to this Outlook and this year's research briefings, but also to earlier efforts. However, given the limited time available, they will focus their remarks on the current Outlook, with the proviso that other topics from previous years can be covered in the discussion to follow.

As this Committee well knows, the Outlooks have had their contentious side, notably in whether they are effective in presenting future problems and opportunities involving science and technology. I personally believe that the previous Outlooks have been extremely valuable in describing the status and likely directions of selected fields of science and technology. Some, however, have argued that the Outlooks have been less successful in highlighting particular policy issues of relevance to the Congressional agenda.

I believe that this report corrects that perceived shortcoming of the Outlook. As this Committee knows, the present format results from a meeting which the Chairman and several members of the Committee, as well as Senator Gorton, had with COSEPUP last year. All of us, of course, will be anxious to have your reaction to the current Outlook, both its contents and the form in which they are presented.

I do, in any case, want to voice my appreciation for the interest and support which the Chairman and, indeed, the full Committee have provided to the scientific and technological endeavor. This hearing is only the latest affirmation of that support, and we are pleased to be part of it.

With that, let me introduce Professor Leon T. Silver, the chairman of COSEPUP. Lee and I are colleagues at several levels. Both of us are geologists. Both of us have taught and done research at Cal Tech, and Lee indeed continues to do so, with his usual vigor and enthusiasm. And both of us care deeply about the health of U.S. science.

Mr. Chairman and members of this Committee, I'm pleased to introduce Dr. Leon Silver.

***

Dr. **Press**. With that, let me then quickly turn to the chairman of COSEPUP, Dr. Lee Silver.

[A biographical sketch of Dr. Silver follows:]

Dr. Leon T. Silver

Present position: Professor of Geology, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125.

Birthdate: April 9, 1925, Monticello, New York.

B.S. (Civil Engineering) 1945, University of Colorado.

M.S. (Geology) 1948, University of New Mexico.

Ph.D. (Geology and Geochemistry) 1955, California Institute of Technology.

1943–46—U.S. Navy.

1945–46—U.S. Naval Civil Engineer Corps (Public Works Officer).

1947–54—Eight field seasons. Field Assistant, Junior Geologist, Assistant Geologist, U.S. Geological Survey, Mineral Deposits Branch, Colorado and Arizona.

1952–55—Research Geologist in Geochemistry of Uranium, AEC Contract Research in the Division of Geological Sciences, California Institute of Technology.

1955–62—Assistant Professor of Geology, California Institute of Technology.

1962–65—Associate Professor of Geology, California Institute of Technology.

1965—Professor of Geology, California Institute of Technology.

1968—Lecturer to NASA Scientists—Astronaut Classes.

1970–76—Geologist, U.S. Geological Survey.

Principal Research Interests.—Igneous and metamorphic petrology; petrology of meteorites; isotope geochemistry of uranium, thorium and lead; application of isotopic techniques in problems of petrology and ore deposits; geochronology; Precambrian history of North America with emphasis on the southwest; batholithic and pre-batholithic rocks of Southern California and Baja California.

Membership and Honors.—Fellow of Geological Society of America and Mineralogical Society of America. Member of American Geophysical Union, Geochemical Society, Meteoritical Society, and Society of Economic Geologists. Fellow of American Association for the Advancement of Science. Honorary Fraternities Sigma Xi, Tau Beta Pi, Phi Kappa Phi, and Sigma Tau. Committee membership in all societies.

Recipient of continuing research grants from NSF and NASA: from AEC 1952–1972, DOE 1979– ; U.S. Geological Survey 1978–1980.

Guggenheim Fellow, 1964.

Chairman, Program Committee, Geochemical Society of America, 1964–1966.

U.S. Representative, IUGS Committee on Establishment of a Worldwide Time Scale, 1968–present.

Recipient of NASA Exceptional Scientific Achievement Medal, 1971.

Recipient of American Institute of Professional Geologists Award for professional excellence, 1972.

Recipient of Geological Society of America certificate of special commendation, 1973.

Member, National Academy of Sciences.

Member, American Academy of Arts and Sciences.

Partipated in the Lunar Exploration and Science Summer Conference held at the University of California, Santa Cruz, California, July 28–August 12, 1967.

Conducted special field geology training for Apollo 13, 14, 15, 16 and 17 crews, September 1969–April 1972. Participated in mission training exercises for all of the crews.

Member of the Lunar Surface Geology Experiment, Apollo 13–17 missions.

Member of the Lunar Sample Preliminary Examination Team, Apollo 15 mission.

Member of the National Aeronautics and Space Administration Science Working Panel, 1971–72.

Member of the National Aeronautics and Space Administration Lunar Sample Analysis Planning Team, 1972–74.

Member of the Universities Space Research Association Advisory Committee for the Lunar Science Institute, 1972–73; 1979– .

Councilor of the Geological Society of America, 1974–76.

Member of the National Aeronautics and Space Administration's Space Program Advisory Council Ad Hoc Subcommittee on Scientist-Astronauts, 1974–June, 1975.

Member of the Board on Mineral and Energy Resources of the National Research Council Commission on Natural Resources, 1975–78; vice-chairman, 1976–77.

Member of the National Air and Space Museum's Visiting Committee, Smithsonian Institution, 1975–present.

Chairman, Working Group for Southwestern and Western U.S., Subcommission on the Precambrian, Commission on Stratigraphy, International Union of Geological Sciences, 1976–present.

Vice-President of the Geological Society of America, 1978.

President of the Geological Society of America, 1979, 1980.

Past President and Councilor of the Geoloical Society of America, 1980.

Chairman, Steering Committee, "Decade of North American Geology", 1980.

Chairman, Committee on Science, Engineering, and Public Policy, National Academy of Sciences. 1984–   ; Member 1979–   .

STATEMENT OF DR. LEON T. SILVER, WILLIAM A. KECK FOUNDATION PROFESSOR OF GEOLOGY, DIVISION OF GEOLOGICAL AND PLANETARY SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CA; AND CHAIRMAN, COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY, NATIONAL ACADEMY OF SCIENCES

Dr. SILVER. Mr. Chairman, Members of Congress, the Committee on Science, Engineering and Public Policy is pleased to present to you *The Outlook for Science and Technology, 1985.* As Dr. Press has stated, we have worked very hard with you, and the product this year, the fourth in the series of *Outlooks* which we have produced, in fact is a significant change from previous products. I would like to hold up for you the first two reports and the new report. [Laughter.]

I believe this is in direct response to your wishes.

The design of the *Outlook* really stems from the continuing interactions with Congress. It is very important to us to have your appraisals of its effectiveness and your suggestions for future improvements, and I hope that the committee will make a point of communicating with us on that.

The current report discusses nine topics on which the committee has conducted briefings. This morning, in a format in which we will attempt to make our formal presentations in a unitized method, we will discuss those briefings and previous briefings.

At this point, we have completed formally 21 briefings in vastly different areas of science, engineering, public health, and we are currently completing a series of briefings which we hope to report on for you at another time, another seven briefings covering a new range of topics.

After we have some presentations from several members of the committee on these briefings, I would like to review for you some of the critical issues which we have identified and for which we believe congressional attention is merited. After that, we will proceed to a period of open discussion which is, I believe, what you want.

Let me, first of all, introduce my colleague, Professor Jacob Bigeleisen. Dr. Bigeleisen is leading professor of chemistry at the State University of New York. Dr. Bigeleisen is going to discuss some of the topics in the physical sciences.

[A biographical sketch of Dr. Bigeleisen follows:]

DR. JACOB BIGELEISEN

Born May 2, 1919, in Paterson, New Jersey. A.B., New York University, 1939. M.S., State College of Washington, 1941, Ph.D. (chemistry) University of California, 1943. Assistant chemist, University of California, 1941–43. Research scientist, SAM Laboratories, Manhattan Project, Columbia University, 1943–45. Research associate, Cryogenic Laboratory, Ohio State University, 1945–46. Fellow, Institute for Nuclear Studies, University of Chicago, 1946–48. Associate chemist—senior chemist, Brookhaven National Laboratory, 1948–68. Professor (chemistry), University of Rochester, 1968–73; department chairman, 1970–75; Tracy H. Harris Professor, 1973–78. Vice President for Research, State University of New York, Stony Brook, 1978–80; Dean, Graduate Studies, 1978–80; Leading Professor of Chemistry, 1978——. Visiting professor, Cornell University, 1953–54. National Science Foundation Fellow and honorary visiting professor, Swiss Federal Institute of Technology, 1926–63. Gilbert N. Lewis Lecturer, 1963. John Simon Guggenheim Fellow, 1974–75. Member: American Academy of Arts and Sciences; American Chemical Society; American Physical Society. Recipient: American Chemical Society Award, 1958; E.O. Lawrence Memorial Award, 1964.

Elected to the National Academy of Sciences in 1966. NAS/NRC activities subsequent to election—Member.—Governing Board of the National Research Council, 1976–80; Executive Committee, Assembly of Mathematical and Physical Sciences [AMPS], 1975, (chairman, 1976–80); Planning Committee for a Natural Resource in Computational Chemistry, [AMPS], 1974–77; Committee on Kinetics of Chemistry Reactions, [AMPS], 1972–76; (member-at-large), Class I Membership Committee, [NAS], 1975; Committee on Bylaws, [NAS], 1970–75; (chairman), Section of Chemistry, [NAS], 1971–74; Executive Committee, Committee on Chemistry and Chemical Technology, [NAS], 1971–74;

Committee on the Chemistry Program, Division of Chemistry and Chemical Technology, 1967.

Nominated for a Position on the Council 1982 Member of Class I.

STATEMENT OF DR. JACOB BIGELEISEN, LEADING PROFESSOR, DEPARTMENT OF CHEMISTRY, STATE UNIVERSITY OF NEW YORK AT STONY BROOK, STONY BROOK, NY; AND MEMBER, COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY, NATIONAL ACADEMY OF SCIENCES

Dr. BIGELEISEN. Mr. Chairman and Members of the Congress, out of the first 21 briefings, there were 8 in the mathematical and physical sciences. I don't want to imply by that that these are the most important topics in science. Rather, they reflect some of the methods by which the physical scientists have organized themselves in this country, namely to get together under the great institution, the National Research Council, and to try to establish what are the important opportunities in science, what are the priorities, and what would it take to do that.

We have had some outstanding reports going back to the early sixties under the auspices of the predecessor committee, COSEPUP, which was organized by my late colleague, George Kistiakowsky. These have been outstanding reports, and one of the most recent ones, the one on astronomy, with which you are familiar, by George Field, has made a great contribution to planning astronomy in the 1980's.

I cannot cover all of the material in all of these briefings, but I will try to give you a little flavor of why this is a time of special opportunity for intellectual advancement in the physical sciences.

We can now study matter and radiation and their interaction with one another on much larger scales than we had been able to just a decade ago. In addition, we can study matter and radiation in greater detail and greater depth.

This recent acceleration of progress stems from the new types of instrumentation available in each of the physical sciences, the use of dedicated microprocessors in the operation of instruments, and in the collection and analysis of data. Finally, access to large high-speed computers allows modeling of systems on a scale which approaches the real world.

But while I stress the great contribution of instrumentation and computers to the physical sciences, our greatest asset is the intellectual capacity of our investigators. I want you to remember that.

Now, I will start and say a few things about some recent advances in physics, the simplest and most basic of the physical sciences. This year, we celebrate the 100th anniversary of the great Danish physicist, Niels Bohr. One can ask, after Bohr and quantum mechanics, what is there left to study and learn about a simple atom in a gas phase?

Recent experiments at the Bureau of Standards use a system of appropriate lasers properly timed and sequenced to slow a beam of sodium atoms, which normally travel at a speed of about 30,000 centimeters a second, to 5 centimeters a second. These atoms are virtually stopped dead in their path. They correspond to a gas within one-millionth of a degree of absolute zero. Such sodium atoms have much sharper energy levels and will certainly improve the accuracy of the measurement of time by atomic clocks. Let me remind you that such clocks are used in navigational satellites.

This is done by lasers, and now I would like to just switch to another application of lasers in the physical sciences, namely chemistry. Lasers have the capacity for us to accelerate and understand chemical reactions at the level of the individual energy states in a molecule. Recent experiments involving a gas, methyl-isocyanide, have shown that you can accelerate that reaction by shining in light from a red laser. If the light from the red laser is absorbed in one of the molecules, the energy is redistributed, the molecule rearranges, and we now have a stable molecule. This experiment has added great insight into our understanding of how a certain fundamental class of reactions, unimolecular reactions, occurs.

Now let me go back to another recent discovery in physics, condensed matter physics. There has been the discovery of the quantum Hall effect. If you take a semiconductor-like silicon and overlay it with a little layer of silicon dioxide and apply a magnetic field, then at the interface, the electrons behave like a two-dimensional gas. They have practically zero electrical resistance, but yet they are not a superconductor.

Now, this new phenomenon is under intense investigation, both in academic and industrial laboratories, because of its great promise for practical applications in types of computers and possibly, also, in the transmission of electrical energy.

Let me say a few words now about some recent advances in chemistry. Catalysis, which is the acceleration of chemical reactions, is important to all types of chemical processing and also to all types of life processes. Catalysts in light processes are enzymes, and they control most of the biochemical reactions in plants and animals. Enzymes are extremely efficient. An enzyme usually is a billion times more efficient than a bulk chemical in promoting a biochemical reaction. Recent advances in the study of materials which are synthesized to imitate enzymes have now approached almost the level of enzymes in catalyzing biochemical reactions, and we can envisage the day when it will be possible to synthesize synthetic materials which will be more efficient and less fragile than natural enzymes.

In the area of astronomy, in the last two decades there have been major discoveries: the quasars, x-ray sources in the cosmos, the pulsars, polyatomic molecules in space, and also the cosmic microwave background, which is a remnant of the big bang.

The Field report has outlined for you the type of major instrumentation needed to maintain the progress in astronomy. In the last issue of *Physics Today*, George Field has given an excellent summary of the positive response by the Congress and the National Science Foundation in maintaining the momentum in that field.

The atmospheric sciences are important in understanding climate, weather prediction, the ozone balance in the stratosphere, and the potential effects of adding gases like carbon dioxide and nitrous oxide to the atmosphere, and such problems as acid rain. The Global Atmospheric Research Program, which was an international program, has provided a very large data base which is currently being used very effectively to improve weather prediction, refine our input data for the calculation of atmospheric circulation, and our understanding of world climate.

Well, with these two great geophysicists to my left, I must now pass to the major advance that has occurred in that field in the last two decades, the concept of plate tectonics. The impact on the field of geological sciences from this concept can hardly be overstressed, and it has provided earth scientists for the first time with a working model of the Earth as a whole. Where the plates slide against one another, as in California, major fault systems develop. Accompanying the interactions of the plate margins are earthquakes, volcanoes, and the concentration of mineral resources.

Knowledge of the stresses in the lithosphere is fundamental to understanding the physics of tectonic processes. Stress measurements at isolated points are the beginning step in the analysis, but it is the spatial variation in stress derived from high-quality digital data from earthquakes with a widespread of geographical distribution and magnitudes that are most important for advances in this field.

To understand the behavior of the Earth at depths greater than 200 kilometers, we have to make measurements at extremely high temperatures and pressures. We now have facilities for the synthesis of large volumes of high-pressure phases in order to determine their thermodynamic and elastic properties. Such measurements can be made with new instrumentation such as lasers, high-pressure diamond anvil cells, and synchrotron radiation.

Now, much of the research I have described requires what we call medium-scale instrumentation. The equipment in a laboratory of a principal investigator involved in this type of work usually amounts to something between $100,000 and $1 million and has an obsolescence time of about 5 years. These certainly create new types of funding problems. In addition, most Federal grants for such instrumentation require significant cost sharing by our institutions, and most of our institutions are now very hard pressed to meet these cost-sharing requirements. In many institutions, principal investigators with great ideas who are being encouraged by the Federal agencies to expand their programs are just told by the Administration: "We won't be able to come up with the matching funds." So this is a new mode that we have to start thinking about. In addition, it is not very useful to buy instruments and not have the money to maintain them and the personnel to operate them, and we have got to pay attention to that.

Now, the high cost of instrumentation in the physical sciences now is having a major impact on bringing new young people into the field. To start up a young investigator, a young faculty member, say, or a young investigator, let us say, in industry costs about a quarter of a million dollars. This now precludes many of our universities from bringing in new young investigators in the physical sciences. Some of them say, "Well, we are just going to recruit established investigators who can bring with them resources of this type." This now, I am concerned, has the potential of drying up the supply of some of the very bright people who would like to enter into research careers and academic careers in the physical sciences.

Thank you.

[The prepared statement of Dr. Bigeleisen follows:]

PREPARED STATEMENT OF JACOB BIGELEISON, MEMBER, COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY, NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF ENGINEERING, INSTITUTE OF MEDICINE

Mr. Chairman, This is a time of special opportunity for intellectual advances in the physical sciences. This is made possible by our increasing ability to study matter and radiation and their interaction with one another on scales much larger, in greater depth and in greater detail than were available just a decade ago. This recent acceleration of progress stems from the new types of instrumentation available in each of the fields, and the use of dedicated microprocessors in the operation of instruments and in the collection and analysis of data. Finally, access to large high speed computers allows modelling of systems on a scale which approach the real world. In my presentation today I will have to restrict myself to a few examples from recent advances in physics, chemistry, astronomy, solid earth sciences, and the atmospheric sciences. I will indicate some of the areas of opportunities for research in these fields in the next five years and some of the national policy issues that have arisen.

Let me start with physics, the most basic of the physical sciences. I will deliberately omit the important and exciting advances and opportunities in high energy and particle physics since you will be holding separate hearings on these subjects. This year we celebrate the 100th anniversary of the birth of Niels Bohr. One can ask, after Bohr and quantum mechanics, what is there left to study or learn about simple isolated atoms in a gas? Recent experiments at the Bureau of Standards use a system of appropriate lasers properly timed and sequenced to slow a beam of sodium atoms, which travel at a speed of 30,000 cm/sec at room temperature, to 5cm/sec. The atoms are virtually stopped dead in their tracks. They correspond to sodium atoms in the gas phase within one millionth of a degree of absolute zero. Such sodium atoms have almost no Doppler broadening, which will certainly improve the accuracy of the measurement of time by atomic clocks. Such clocks are used in navigational satellites.

Lasers applied to chemical systems have the capacity to accelerate chemical reactions and provide new depths of understanding about how chemical reactions occur. For instance the reaction in which methyl-isocyanide $H_3C\text{-}NC$ is converted to the more stable species methyl-cyanide $H_3C\text{-}CN$ does not occur in the gas phase at an app.eciable rate until the substance is heated to a few hundred degrees centigrade. The reaction proceeds rapidly if the methyl-isocyanide is illuminated with a red laser. The laser energy is absorbed by the C-H stretching vibrations in the molecule and is then transferred within the molecule to the motion between the methylcarbon and the nitrogen atom. This provides the energy for the molecule to rearrange and form methylcyanide. The study presents a remarkable confirmation of the (RRKM) theory of unimolecular reactions.

" Let me return to a more complex problem in physics, condensed matter physics. The recent discovery of the quantum Hall effect shows that in a material like silicon overlayed by silicon oxide, the electrical resistance at the interface in a magnetic field is virtually zero, yet the material is not a superconductor. The temperature coefficient of the electrical conductivity is very small. The electrons at the silicon-silicon dioxide interface behave like a two dimensional gas. These systems are currently under intense investigation because of their intrinsic scientific interest as well as their potential applications in supercomputers and power transmission.

In metallic compounds, the "effective" mass of an electron is of the order of ten times the mass of a free electron as a result of coupling between the electrons and the atomic lattice. Recent experiments with stoichiometric compounds of uranium and beryllium give "effective" electron masses some 100 to 1000 times the free electron mass. None of our current theories can explain these results.

Catalysts are important in promoting chemical reactions. Enzymes are a particularly important set of catalysts since they control most biochemical processes in plants and animals. They are extremely efficient. For instance enzymes which hydrolyze esters are a billion times more efficient than bulk chemical catalysts, such as hydrogen ions, for the hydrolysis of esters. Recent work on the synthesis of biomimetic catalysts, compounds prepared and structured to help us understand how enzymes function, now approach enzymes such as chymotrypsin in efficiency. We may be approaching the day when we will be able to synthesize catalysts for biochemical processes which are more efficient and less fragile than enzymes.

Two decades ago the synthesis of organic compounds and new drugs was the work of a chemical artist, who proceeded by some general rules and brilliant intuition. The latter has now been augmented by the deep understandings of chemical reactivity derived from theory and experiment. Much of this information has now been systematized on computers with the result that the computer has become an important tool for strategy planning in the synthesis of new organic compounds and drugs from simple building blocks.

Our knowledge and understanding of the universe has expanded in the last two decades by the discoveries of quasars, x-ray sources in the cosmos, pulsars, polyatomic molecules in interstellar space and the cosmic microwave background radiation, a remnant of the big bang. Space telescope, and a number of new facilities, including an advanced x-ray astrophysical facility, radio telescopes, the new large optical telescope being planned by the California Institute of Technology and the University of California, and infrared telescopes in space will greatly expand our observational opportunities in all regions of the electromagnetic spectrum.

Atmospheric sciences are important for understanding climate, weather prediction, the ozone balance in the stratosphere, the potential thermal effects resulting from the addition of carbon dioxide, nitrous oxide and other gases to the atmosphere, and such environmental problems as acid rain. The Global Atmospheric Research Program (GARP) has provided a large data base which is currently being used effectively to improve weather prediction, refine our input data for the calculation of atmospheric circulation, and our understanding of world climate.

Uncertainties in our understanding of atmospheric chemistry necessary for understanding the ozone balance have been significantly reduced by laboratory experiments and laser analysis of the composition of the upper atmosphere and the stratosphere. Access to supercomputers on a full-time basis is having a major impact on the accuracy of weather prediction.

The impact of the concept of plate tectonics, introduced twenty years ago, on the geological sciences can hardly be overstressed. It provided earth scientists for the first time with a working model of the earth as a whole. Where the plates slide against each other - as in California - major fault systems develop. Accompanying the interactions of the plate margins are earthquakes, volcanoes, and the concentration of mineral resources.

With the remarkable accuracy of new Global Positioning Satellite receivers, we can foresee that within a few years plate velocity vectors can be determined and accelerations in plate motion can be measured. The interaction of the plates in great earthquakes can be observed directly by measurement of earthquake displacements and post-seismic dissipation of strain.

Knowledge of the stresses in the lithosphere, the earth's rigid outer shell, is fundamental to understanding the physics of tectonic processes. Stress measurements at isolated points are the beginning step in analysis, but it is the spatial variation in stress derived from high quality digital data from earthquakes with a wide spread of geographical distribution and magnitudes that are most important.

At levels in the Earth's upper mantle (depths greater than 200 kilometers), realistic interpretations require measurements of the physical properties of individual minerals and aggregates at temperatures and pressures relevant to the mantle. Elastic constants and their pressure and temperature derivatives, rheological properties, and phase relations are vital to these considerations. Facilities for synthesis of large volumes of high-pressure phases in order to determine thermodynamic and elastic properties are essential. Such measurements can now be made with the new instrumentation that has become available, e.g. lasers, high pressure diamond anvil cells and synchrotron radiation.

Research of the type I have described requires medium scale instrumentation. The equipment in a laboratory of a principal investigator requires an investment between $100,000 to $1,000,000. Astronomers, in addition, need access to large scale national facilities; physicists and chemists require facilities such as dedicated synchrotrons and high intensity neutron sources; while the atmospheric scientists and geoscientists participate in national and international cooperative programs. Purchases of medium-scale instruments inevitably entail special budgetary procedures, in many cases involving pooling mechanisms such as institutional and regional facilities. Further, most federal grants for medium-scale instrumentation require significant cost-sharing by institutions. Realistically, however, institutional resources, even those of the major research universities, are unable to meet all of the requests from their faculties for cost-sharing.

Grants to buy instruments need to be accompanied by funds for operation and maintenance throughout their expected lifetimes. The obsolesence time of much current medium scale instrumentation is about five years. Funds to purchase instruments often do not include money to operate them or to hire and train technical personnel to help investigators use state-of-the-art technology. Sometimes, it is easier for academic institutions to obtain financing for new equipment than for operating and upgrading existing equipment, even though the latter may be only a year old. Therefore, it would seem prudent for granting agencies to consider policies for funding capital equipment that allow money for maintenance, operation, upgrading, and training or hiring technical personnel.

In any case, concern for the inadequacy of advanced instrumentation for academic research is likely to increase in the next few years. There is a significant gap in access to medium-scale, technologically advanced instrumentation between academic and industrial researchers working in the same area. While several federal agencies -- such as the National Science Foundation, the Department of Defense, and the Department of Energy -- have initiated instrumentation programs to bring state-of-the-art equipment into university laboratories, these programs tend to be small compared to the magnitude of the problem. It is estimated that the current backlog for medium scale instrumentation in the fields discussed in this satatement exceeds one billion dollars.

An even more severe situation exists with respect to major facilities because of the costs, the lack of any significant programs for major facilities other than instruments, e.g., large astronomical observatories, and the virtual absence of a process by which comprehensive merit review can be utilized to allocate resources for new facilities, which including renovated and new buildings for scientific research.

The cost of equipment for a new investigator to embark on a research career in the physical sciences is of the order of $200,000. These funds must come from institutional resources. This precludes many universities from making appointments. Even those research universities with the greatest financial resources are finding it difficult to meet these costs. The result is a pattern of shifting away from bringing young investigators into the system in favor of attracting established investigators who are better able to bring external resources with them. Additional programs, complementing the Presidential Young Investigators Awards, are necessary to maintain the flow of bright young talent into academic careers in the physical sciences.

Dr. SILVER. Thank you, Dr. Bigeleisen.

I would now like to introduce Dr. Floyd Bloom, Director and Member of the Division of Preclinical Neuroscience and Endocrinology at the Scripps Clinic, La Jolla.

Floyd.

[A biographical sketch of Dr. Bloom follows:]

DR. FLOYD E. BLOOM

Date and Place of Birth: October 8, 1936; Minneapolis, Minnesota.

Citizenship: United States.

Marital Status: Two children; Married Jody P. Corey.

Education: 1956—A.B., cum laude, Southern Methodist University, Dallas, Texas. 1960—M.D., cum laude, Washington Univ. Sch. Med., St. Louis, Missouri.

Professional training.—December 1983—present: Director and Member, Division of Preclinical Neuroscience and Endocrinology, Scripps Clinic and Research Foundation, LaJolla, California.

October 1975——.—Director, Arthur V. Davis Center for Behavioral Neurobiology, The Salk Institute, San Diego, California.

July, 1973 to Oct., 1975: Acting Director, Division of Special Mental Health Research Programs, NIMH, St. Elizabeths Hospital, Washington, D.C.

July, 1969 to Oct., 1975: Chief, Laboratory of Neuropharmacology, Division of Special Mental Health Research Programs, NIMH, St. Elizabeths Hospital, Washington, D.C.

Aug., 1968 to July, 1969: Acting Chief, Lab. of Neuropharmacology, Division of Special Mental Health Research Programs, HIMH, St. Elizabeths Hospital, Washington, D.C.

July, 1968 to Oct., 1975: Chief, Section on Cytochemical Pharmacology, Lab. Neuropharmacology, Division of Special Mental Health Research Programs, NIMH, St. Elizabeths Hospital, Washington, D.C.

July, 1964 to July, 1968: Yale Univ. Sch. Med; USPHS Special Post-doctoral Fellow in Anatomy, 1964–1966; Assistant Professor Anatomy (Pharmacology), 1966; Assistant Professor Pharmacology (Anatomy), 1967; Associate Professor Pharmacology, 1968; Research Career Development Awardee, NINDB, 1966-1968; Faculty Advisor to Biological Sciences Training Program, Dept. Psychiatry, Connecticut Mental Health Center, 1966-1968.

July, 1962 to July, 1964: Research Associate, NIMH at St. Elizabeths Hospital, Clinical Neuropharmacology Research Center; Associate in Physiology, George Washington Med. Sch., Washington, D.C.

July, 1960 to June, 1962: Intern and First Year Assistant Resident, Ward Medical Service, Barnes Hospital, St. Louis, and Assistant in Medicine, Washington Univ. Sch. Med. Fall, 1959: Assistant Instructor, Dept. Pharmacology, Washington Univ. Sch. Med.

Military Service: USPHS—1962-1964, Surgeon (Lt. Cmdr.)

Societies.—Phi Beta Kappa, Alpha Omega Alpha, Sigma XI, New York Academy of Science, American Society for Cell Biology, American Association of Anatomists, American Society for Pharmacology and Experimental Therapeutics, American College of Neuropsychopharmacology, Society for Neuro-science, Secretary 1973-1975, President 1976-1977, Psychiatry Research Society, International Society for Sterology, American Physiological Society.

Honors and Awards.—A.E. Bennett Award, 1971; A Cressy Morrison Award, 1971; Arthur S. Flemming Award, 1972; Mathilde Solowey Award, 1973; Biological Sciences Award, 1974; Washington Academy of Arts and Sciences; Member, National Academy of Sciences, 1977; Member, National Advisory Mental Health Council, 1976-1980; Member, Research Work Group, President's Commission on Mental Health, 1977; Adjunct Professor, Department of Neurosciences, University of California, San Diego, 1977; Member, Alan T. Waterman Award Committee, NSF, 1978-1980; Member, Richard Lounsbery Award Jury, NAS, 1979-81; Chairman, Neurobiology Section, NAS, 1979-83; Distinguished Alumnus Award, Washington University, 1980; National Assoc. of Mental Health McAplin Award Recipient, 1980; Member, National Commission on Alcoholism and Other Alcohol-Related Diseases.

STATEMENT OF DR. FLOYD E. BLOOM, DIRECTOR AND MEMBER, DIVISION OF PRECLINICAL NEUROSCIENCE AND ENDOCRINOLOGY, SCRIPPS CLINIC AND RESEARCH FOUNDATION, LA JOLLA, CA; AND MEMBER, COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY, NATIONAL ACADEMY OF SCIENCES

Dr. BLOOM. Mr. Chairman, it is a pleasure to be back with you again.

When Dr. Low and I were here for the previous *Outlook* report, we had an exciting morning and talked about a large number of areas of science.

This morning, because of the wealth of talent on our panel, I am going to concentrate on the life sciences and lay out a few of the points that the current *Outlook* report establishes from the examples in the life sciences: oncogenes, the biology of the cells of the walls of the blood vessels, and parasitology, parasitism.

This is a very exciting time in biology because even since 1984, tremendous advances have continued to flow from a variety of scientific disciplines working together to help improve our understanding of the way the cells of the body work together and the molecules that allow certain kinds of cells to work in unique ways.

When we talk about the causes of cancer, we now do so with a great deal of wisdom, whereas before, we had a great deal of curiosity and stress. Now, because of the advances in the field that we might call genetic engineering, the biology of molecules that control the expression of our genetic information, it is possible to understand the various levels at which environmental and infectious agents can cause cancer.

Even in this morning's paper was a thrilling example, through death, of understanding the relationship between the biological systems in the body and what we call cancer. David, the little boy who lived in the bubble because he lacked the appropriate kinds of immune defenses, when he received bone marrow that might have saved his life, died because the viruses in the bone marrow cells he took from his sister produced a cancer that later spread all over his body and killed him.

Now, what that illustrates to us is that there can be many ways to cause cancer; that the immune system of the body, constantly on guard against the causes of cancer, can often protect us; but, under certain conditions, it is not sufficient to do so.

The molecules that are called oncogenes are genes that express molecules in cells and turn them into cancerous cells. Those genes are present in our cells all the time but are often triggered very inadvertently by environmental and other kinds of challenges to the body.

The lessons from the life sciences really are that the more we understand about the molecules that exist in normal cells, the more we understand about the ways that diseases can be caused, and the more we then understand about the ways that health can be maintained and perhaps diseases be prevented in the future. So, this area continues to be one in which the life sciences, because of the ability to understand the chemistry of the cells—what makes them similar and what makes them different—continues to be one of tremendous importance to the country because it covers so many aspects of life.

Not only is it appropriate to deal with life sciences under the topic of biology, health, and medicine, but also in terms of agriculture, one of our country's strong points. The ability to understand how the cells of plants express their genetic information, why certain plants can grow under certain conditions and why certain plants die in certain kinds of soils, is all potentially explainable by the study of the expression of genetic information, and we believe that these areas offer great importance for future study and ought to be maintained in as vigorous a fashion as possible. Although there have been discussions of doing fundamental research in agriculture, it has been difficult to bring that off because of many of the same problems that Dr. Bigeleisen described for the physical sciences.

There are two sets of issues that we proposed in the report that Congress might wish to direct their attention to specifically. They have to do with the use of genetic engineering and the possibility that this needs to be applied to a variety of aspects of our life but must be applied in a way that maintains full public understanding of what the risks and the opportunities are. The existing monitoring mechanisms have proven to be very effective, but one, of course, would like to be sure that those are implemented and continued.

It is important, also, to understand about human biological engineering. We read a lot about this in the papers from time to time, about the ability to repair genes and to insert genes into cells. This technology has indeed moved a great deal, even in the last 2 years, but the ability to put a gene that you would like into a cell and have that cell use it in a proper way is far from an everyday experience, and so tremendous amounts of research are going to be necessary in order to bring this exciting discipline really to application for human beings.

I would just like to conclude by saying that the life sciences and the physical sciences share some of these general phenomena that influence the course of scientific development; that is to say that life scientists have a problem in instrumentation. As the field has evolved and as the methods have improved, the instruments that one does research with have changed, and so it is very difficult for new scientists in the life sciences to get a start without a large amount of equipment money to be able to initiate their laboratory settings. This means that certain places can concentrate, whereas others will not be able to bring their life sciences faculty up to par with the rest of the U.S. research effort.

In terms of manpower, we need new hands and new minds as well, and it is important to keep them coming in. But right now, there is a wealth of attraction that brings young scientists into the life sciences and a dearth of jobs where they can continue to apply what they have learned during their training period.

In the field of health medicine, which is one that concerns me greatly, we are concerned about a missing training for two kinds of people. One is basic scientists who are aware enough of clinical problems to be able to turn their attention to the fundamental issues of the clinical problems. That is why my laboratory, for example, is called preclinical, because we try to focus on those basic scientific developments that allow us to attack clinical problems.

The other kind of scientist that we are worried about is the scientifically-trained clinician. As the opportunities for clinical training have decreased in the country, the opportunity to do research while doing your clinical training has decreased even more, and without scientifically-trained doctors, we will not be able to make observations in the scientific way that is necessary to gain insight into the nature of disease.

Let me just stop by saying that biology still has excitement, and I am sure, when COSEPUP appears before you in its next appearance here, that there will be additional examples from the field of biological research.

We hope that those examples will come from science being done in the United States, because this is a highly competitive area, and if we lose our momentum, those discoveries will be made elsewhere, and that would not be good for us educationally, intellectually, or economically.

Thank you, sir.

[The prepared statement of Dr. Bloom follows:]

PREPARED STATEMENT OF FLOYD E. BLOOM, MEMBER, COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY, NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF ENGINEERING, INSTITUTE OF MEDICINE

Mr. Chairman, It is my pleasure to participate in this joint presentation of the issues arising from the preparation of the 1985 Outlook Report on Science and Technology. I chaired the Study Group that developed the Third Outlook which we discussed before this Committee in January of 1984. And I was also a member of the Sub-Committee that produced the volume we are discussing today. The Chairman of our Committee on Science, Engineering and Public Policy (COSEPUP), Dr. Leon Silver, has asked me to focus in my testimony on issues concerning the life sciences.

To begin: we should be aware of the extremely rapid progress within the life sciences, with concepts and fundamental principles undergoing major change as new scientific understanding accumulates at an unprecedented rate. Thus, it is important that a continuing dialog be maintained to communicate widely the implications and opportunities of these scientific advances.

The decade of the 1980's has been recognized as the era of biology. Our understanding of life processes has been clarified enough that man can begin to understand how specific molecules function within specific cells to maintain health and to understand, in part, what happens to all functions in disease. The key point is that a strong and fundamental understanding is needed to attack effectively the major illnesses of man. An example is AIDS. Without the fundamental understanding that has accumulated over the past 2 decades on the immune system of the body, and the extensive characterization of the antigen-recognizing cells, the antibody-producing cells and the chemical signals by which they work together to mount an immune response, (see the 1984 COSEPUP Research Briefing on Immunology) it would not have been possible to understand the nature of the Acquired Immune Deficiency Syndrome, or AIDS, within the short period after it was first recognized. Major health problems -- regardless of their costs -- are not immediately amenable to effective scientific understanding until this scientific base has been developed.

Three further general points deserve specific attention within our consideration today of the implications to be drawn from the life sciences.

One point deals with how cells function. As the life sciences have gained momentum over the past decade, the principal advances have been in clarifying how the genetic information contained within every cell of the body achieves its unique functional capacity to act as a liver cell, a kidney cell, a muscle cell, or a nerve cell. The more that has been learned about the shared and unique aspects of cellular function in terms of which genes within any particular type of cell are expressed, the better we understand the sites at which disease may alter normal cell function. Furthermore, similarities among cells may begin to overshadow their differences. This is important because many common factors share the ability to regulate the activity of different cells by acting through the blood stream. (These factors include, for example, the hormones of the endocrine system described in the previous Outlook report.) Similar principles of cellular regulation exist throughout the body; and, indeed across almost all forms of animal life, and many forms of plant life.

Thus, we can see that the lessons learned from various life forms and their specific disorders can have very wide ramifications: information on how viruses reproduce in cells has led to valuable insights into what may cause cancer; information on how the immune system distinguishes foreign proteins from self-made proteins has provided clues to how certain infectious organisms -- whether they be neuron-seeking viruses or the red blood cell-seeking parasites of malaria to identify and invade their preferred target cells. The same research has provided important insight into organ transplantation therapy. The lessons to be learned and applied are broad. A critical need now is to apply general findings to specific cellular systems involved in human diseases.

Genetic Engineering:

My next point deals with genetic engineering. As indicated in this Outlook, the powerful research tools that have emerged from the scientific study of cellular genetics have not simply accelerated the rate of discovery in the life sciences. These same tools have also already yielded promising therapeutic applications in clinical medicine, veterinary medicine, agriculture, energy and pollution control. The 1985 Outlook outlines the processes by which the federal government has coordinated research involving genetic engineering. The issue is now the wider commercialization of these techniques. To that end, the Cabinet Council Working Group in Biotechnology has developed a framework proposal for the coordinated regulation of biotechnology. Against that background, the 1985 Outlook poses two issues for the Congress:

1) how to maintain oversight of what have been remarkably effective mechanisms for monitoring the use of recombinant DNA and related technologies;
2) how to assure a full, but balanced, consideration of the requirements for public safety and the needs and opportunities of an emerging industry of high biological technology.

My third point is related to the previous one, but deals particularly with human biology. The advances in understanding the molecular mechanisms underlying genetic expression and regulation have been enormous. However, the capability to apply this knowledge through applied engineering of human and animal genes is far from clear. Although many genes can be isolated, and the basis for inheritable disorders has in several cases been identified, it is not now possible to replace these defective genes in a embryonic human. The reason is that it is impossible to be certain that the proper genes get into the correct cells. It is even more problematic to get them into the correct place within the wealth of genetic information inside every cell. Thus, it is not yet possible to assure the proper expression of a gene even if that gene can be isolated or synthesized and injected into the proper population of cells.

Further, in discussing genetic engineering of cells, especially human cells, it is critical to distinguish between manipulating genetic information in the germline cells, by which parents can transfer their genetic information to their offspring, from the far less difficult transfer of genetic information into specific types of malfunctioning somatic cells, such as bone marrow "stem" cells that can seed the production of healthy red or white blood cells.

Let me briefly turn from a prospective issue in human biology to one that already exists. Presently, new biological methods make possible the fertilization of human eggs outside of the body, with or without the opportunity for exposure to possible corrective genetic manipulations, followed by implantation to a maternal host. While this methodology is in use in some parts of the world, it is not an area that can be studied scientifically in the United States, owing to a de facto ban. There is also relatively little work possible now on improving the methods by which genetic information may be introduced into human somatic or germline cells. Without this information, insertion of genetic material into germline cells or even into the rapidly turning over cells of the bone marrow will remain an unpredictable art. Rather than offering potential repair of fatal genetic defects, such efforts may produce still other deleterious effects.

Overall, there are two issues within the province of human biology worthy of Congressional attention. One is for the Congress to continue its oversight through the NIH of plans for human gene therapy. The second is to encourage research programs to enlarge our fundamental understanding of fetal biology.

Future Concerns in Life Sciences Research:

Let me conclude by noting for the record that the life sciences share common problems with the other sciences. These include maintaining both an adequate influx of new minds and hands and the capacity of the research setting as new, more powerful instruments are developed and as the existing research environments decay. A particular personnel issue unique to the life sciences is that of a major gap in the application of modern biological research to pressing human health problems owing to the loss of critical personnel in the interface between the basic sciences and the clinical sciences.

Congress has been presented with series of plans to help retain the active clinically trained practioner in a research environment. But it is not at all clear that any of these plans are succeeding in restoring the former balance of medically trained personnel in the biomedical sciences. Two types of researchers may be needed: 1) clinically-exposed scientists who have sufficient insight into the mechanisms of disease to be able to apply novel biological solutions; and 2) scientific clinicians who can bring to the bed side and the clinic the ability to ask heuristic questions about the unsolved clinical problems that remain important challenges to the health of our citizens.

Finally, I wish to point out that the possibilities opened by the magnificent achievements in biology will remain just that without the continued nourishment provided by Congress over the past decades. To be forced to call a recess at this moment of expectant application would be counterproductive, intellectually, societally. Such a recess, if forced by the momentary difficulties of the national economy, would also potentially translate into a series of lost opportunities and lives, and the likely possibility that the important discoveries and their applications and economic gains will be won elsewhere.

Thank you.

Dr. **Silver**. Thank you, Floyd.

I would now like to introduce Dr. Edward Mason, with a newly formed company called Amoco. [Laughter.]

He is the vice president for research for Amoco. Dr. Mason is going to describe for us some of the important new activities in the areas of engineering and applied science.

[A biographical sketch of Dr. Mason follows:]

Dr. Edward A. Mason

Office.—Vice President, Research, Amoco Corporation, P.O. Box 400, Naperville, Illinois 60566, Tel: 312/420-5252.

Home.—145 Hillcrest Avenue, Hinsdale, Illinois 60521, Tel: 312/920-0664.

Birth.—August 9, 1924; Rochester, New York.

Education.—B.S., Chemical Engineering, University of Rochester, 1945; S.M., Chemical Engineering Practice, M.I.T., 1948; Sc.D., Chemical Engineering, M.I.T., 1950.

Work experience.—1950-1953—Department of Chemical Engineering, M.I.T.; Assistant Professor (Instructor, 1949); 1953-1956—Director of Research, Ionics, Incorporated; 1957-1975—Departments of Chemical and Nuclear Engineering, M.I.T.; 1957—Associate Professor; 1957—Senior Design Engineer, Oak Ridge National—Laboratory (on leave from M.I.T.); 1963—Professor; 1965-1966—Senior Postdoctoral Fellow National Science Foundation Research at Euration Research Center, Ispra, Italy (on leave from M.I.T.); 1967—Director, Agro-Industrial Energy Project, Oak Ridge National Laboratory (on leave from M.I.T.); 1971-1975—Professor and Head, Department of Nuclear Engineering.

1975-1977—Commissioner, U.S. Nuclear Regulatory Commission (on leave from M.I.T.).

1977—Vice President, Research, Standard Oil Co. (Indiana).

Consulting experience.—1957-1975—Chemical and nuclear engineering consultant to several Government agencies and laboratories and industrial concerns in electrochemistry, electrodialysis, ion exchange, solvent extraction, fuel cells, water desalting, fluidized bed technololgy, nuclear reactor and nuclear fuel processing design, safety and economics, Consultant to several foreign governments on nuclear energy. Member, U.S. AEC Advisory Committee on Reactor Safeguards 1972-75.

Honorary and professional society memberships.—National Academy of Engineering (Council 1978-84); American Academy of Arts and Sciences (Fellow), New York Academy of Sciences, American Institute of Chemists (Fellow), American Nuclear Society (Fellow), American Chemical Society, American Institute of Chemical Engineers, American Petroleum Institute, American Association for the Advancement of Science (Fellow), Western Society of Engineers (Board of Directors 1979-85), Industrial Research Institute (Company Representative), Directors of Industrial Research (Member), Phi Beta Kappa, Sigma Xi, Tau Beta Pi.

Recipient, Robert E. Wilson Award, American Institute of Chemical Engineers, 1978.

Registered Professional Engineer (Massachusetts).

Boards and advisory committees.—Member, M.I.T. Corporation Development Committee (1982-1985); Member, M.I.T. Corporation Visiting Committee for Department of Mechanical Engineering (1982-1985); Member, M.I.T. Advisory Board for the Center for Advanced Engineering Study (1981-1984); Member, University of Chicago Visiting Committee, Division of the Physical Sciences (1977-85); Member, National Advisory Board, Georgia Institute of Techology (1981-1987); Member, Board of Directors, The John Crerar Library (Treasurer) 1977-; Member, Board of Directors, Robert Crown Center for Health Education, 1977-; Cetus Corporation, Member, Board of Directors (1977-84); Member, Board of Directors, Commonwealth Edison Company, 1980-; Member, Board of Directors, XMR, Inc. (1984-); Member, National Research Council Committees: Committee on Science, Engineering and Public Policy (1983-87), Energy Engineering Board (1983-85), Steering Committee NAS/NAE China Initiative on Technology and Industrial Development.

Publications.—Author of over 75 publications, 4 patents.

Family.—Wife, Barbara Earley Mason; 6 children.

STATEMENT OF DR. EDWARD A. MASON, VICE PRESIDENT, RESEARCH, AMOCO CORP., NAPERVILLE, IL; AND MEMBER, COMMITTEE ON SCIENCE ENGINEERING, AND PUBLIC POLICY OF THE NATIONAL ACADEMY OF SCIENCES

Dr. MASON. Thank you, Lee.

Mr. Chairman and members of the committee, I am pleased to have been invited to join my scientific colleagues here and look forward to joining in the general discussion.

In my brief testimony, I shall concentrate on four technical areas which relate closely to engineering applications of fundamental research in both science and engineering. They are: materials science; the related area of advanced polymeric composites; chemical and process engineering for biotechnology; and computers in design and manufacturing.

Among the past accomplishments in the fields of materials science and advanced composites is the exploration of new materials that may have unusual or unexpected combinations of properties and which have the potential for reducing U.S. dependence on foreign sources of strategic materials. Two examples are dislocation-free silicon, the building block of silicon-chip technology, and also fiber-reinforced composites, used in very strong, light-weight structures for airplanes and increasingly in automobiles.

While the United States has its own position in materials science and advanced polymeric composites, vigorous programs are also proceeding in Europe and Japan. Research in materials science and advanced polymeric composites calls for communication and collaboration among a number of different disciplines. These groups tend to be insular in the present U.S. university structure. The future will belong to those who can bring together and support the various necessary disciplines in order to construct a coherent intellectual structure that will form the basis for understanding present achievements and for advancing to new levels.

The phenomenal progress in molecular biology, genetics, and biochemistry in the past 20 years, which Dr. Bloom has just referred to, now makes it possible to have cells manufacture products ranging from simple molecules to complex proteins. Potential opportunities for applying biochemical technologies are diverse.

In the area of human and animal health care, for example, a new family of products based on genetically-engineered proteins may emerge that can detect quickly and accurately viral and bacterial diseases, susceptibility to autoimmune disease, genetic defects, and neoplasms.

In agriculture, the new technologies may yield improved crop plant and animal species. Other prospects lie in environmental protection, where biochemical engineering may provide methods of destroying or removing toxic products.

The United States has a very strong capacity for leadership in the businesses based on biotechnology, owing largely to the preeminence of the basic biology research conducted in American laboratories. However, leadership in commercial exploitation will not be achieved without adequate knowledge of the fundamental engineering principles underlying bioprocess scale-up. West Germany, Japan, and Great Britain already have mounted national programs to develop strong biochemical industries.

Issues for congressional consideration are: strengthening the link between life scientists and biochemical engineers, both in the training of graduate students as well as in incentives for quality faculty to dedicate their careers to this area; also, funds to enable academic units to purchase the equipment essential to contemporary research; and support of cross-disciplinary and university-industry cooperative research.

The rapid growth of computer technology and the promise of improved productivity and quality have drastically changed the way we can design and manufacture products. We can think of design and manufacturing, together with marketing, as three inseparable parts of our production cycle from conception to delivery.

Marketing provides information to engineering and design on the product required. Engineering designs the product and transfers product-definition data to manufacturing. As an example of the usefulness of computer-aided design, designers at General Motors were able to reduce the total weight of an average automobile by 25 percent and, at the same time, improve the overall structural integrity of the car.

The central issue in the field of computers in design and manufacturing is a lack of scientific knowledge. This lack of knowledge is less pervasive than it was some 10 years ago, in part because university research has made some significant accomplishments with very limited resources. The large integration efforts mentioned in the *Research Briefings* will require a much stronger base of fundamental knowledge than is now available.

In summary, the areas I have discussed have all grown rapidly in the recent past and offer great opportunities for maintaining and improving the industrial and economic competitiveness of the United States. They are high-payoff areas with clear opportunities for both short- and long-range applications.

Much of the underlying knowledge on which these fields are based is the result of Federal funding of basic science. They represent good examples of the large benefits which the public can realize from farsighted, long-range funding of basic research in both science and engineering.

The enormous development of new products and markets that have resulted from the process of transfer of fundamental scientific findings into commercial use by the application of advanced engineering is well known to all of us, and our Nation's competitive strength in science and technology must continue to be fostered in the future.

Thank you.

[The prepared statement of Dr. Mason follows:]

PREPARED STATEMENT OF EDWARD A. MASON, MEMBER, COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY, NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF ENGINEERING, INSTITUTE OF MEDICINE

Mr. Chairman, I am pleased to have been invited to join my scientific colleagues in testifying here and look forward to joining in the general discussion.

For my brief testimony, I shall concentrate on four technical areas which have been the subject of Research Briefings prepared under the auspices of COSEPUP in recent years and which relate closely to engineering applications of fundamental research in both science and engineering. They are:

(1) Materials Science (1982), and the related area of
(2) Advanced Polymeric Composites (1984)
(3) Chemical and Process Engineering for Biotechnology (1984)
(4) Computers Design in Manufacturing (1983)

Among the past accomplishments of this type of inderdisciplinary research in the fields of materials science and advanced composites is dislocation-free silicon--the building block of silicon-chip technology and fiber-reinforced composites used in very strong light-weight structures for airplanes and increasing in automobiles. Another important activity within materials science is the exploration of new materials that may have unusual or unexpected combinations of properties, and which have the potential for reducing U.S. dependence on foreign sources of strategic materials.

The phenomenal progress in molecular biology, genetics, and biochemistry in the last 20 years now makes it possible to have cells manufacture products ranging from simple molecules to complex proteins. The United States leads the world in research in the science of the "new" biology. Potential opportunities for applying biochemical technologies are diverse and provocative. In the area of human and animal health care, for example, a new family of products based on genetically engineered proteins may emerge that can detect quickly and accurately viral and bacterial diseases, susceptibility to autoimmune diseases, genetic defects, and neoplasms. Other proteins, such as those inhibiting the growth of tumors or those that dissolve blood clots, are being tested. In agriculture, the new technologies may yield improved crop plant and animal species, and fungicides and herbicides that are highly potent, specific, and environmentally safe. Other prospects lie in environmental protection, where biochemcial engineering may provide methods of destroying or removing toxic products.

Although the use of computers in design and manufacturing is relatively new, the rapid growth of computer technology and the promise of improved productivity and quality have drastically changed the way we can design and manufacture products. At the risk of oversimplifying, we can think of design and manufacturing together with marketing as three inseparable parts of our production cycle from conception to delivery. Marketing provides information to engineering and design on the product required. Engineering designs the product and transfers product-definition data to manufacturing. As an example of the usefulness of computer aided design (CAD), designers at General Motors were able to reduce the total weight of an average automobile by 25% and at the same time improve the overall structural integrity of the car. CAD reduced the time required to solve complex stress equations, allowing designers to increase the number of iterations and evaluate more completely the structural changes when metal was removed.

The United States has a strong capacity for leadership in businesses based on biotechnology, owing largely to the basic research conducted in American laboratories. A tremendous number of opportunities await commercial exploitation. This will not occur rapidly without adequate knowledge of the fundamental engineering principles underlying bioprocess scale-up. West Germany, Japan and Great Britain already have mounted national programs to develop strong biochemical industries, achieving industrial leadership in biotechnology for the United States requires a wider knowledge base than is now available, a greater number of trained personnel, support of pilot studies for biochemical engineering processes, and working connections between basic biological research and engineering practice.

Issues for Congressional consideration are: strengthening the links between life scientists and biochemical engineers, both in the training of graduate students, as well as in incentives for quality faculty to dedicate their careers to this area; funds to enable academic units to purchase the equipment essential for contemporary research; and support of cross-disciplinary and university-industry cooperative research.

While the United States has its own position in materials science and advanced polymeric composites, vigorous programs are also proceeding in Japan and Europe. The science of materials processing would be greatly advanced by increased understanding of some basic phenomenon that pervade many processing techniques. Research in materials science and advanced polymeric composites calls for communication and collaboration among several disciplines, including chemistry, physics, chemical and mechanical engineering, materials science, computer science, and toxicology. These groups tend to be insular in the present U.S. university structure. The future will belong to those who can bring together and support the various necessary disciplines, in order to construct a coherent intellectual structure that will form the basis for understanding present achievements and for advancing to new levels.

Universities, industry, and government traditionally have been the key institutions in research on manufacturing technology. A national effort bringing together these three segments of our technical community is required if we are to compete effectively in worldwide markets. The central issue in the field of computers and design and manufacturing is a lack of scientific knowledge. This lack of knowledge is less pervasive than it was ten years ago in part because university research has made some significant accomplishments with very limited resources. The large integration efforts mentioned in the research briefings will require a much stronger base of fundamental knowledge than is now available.

In summary, the areas I have discussed have all grown rapidly in the recent past and offer great opportunities for maintaining and improving the industrial and economic competitiveness of the United States. They are high payoff areas with clear opportunities for both short and long range applications. Like many other areas of science and technology, they have been characterized lately by very rapid progress resulting from the combined efforts of several fields of science and engineering. Much of the underlying knowledge on which these fields are based is the result of federal funding of basic science. They represent good examples of the large benefits which the public can realize from farsighted, long range funding of basic research in both science and engineering. This point was clearly stated in the recent report by the President's Commission on Industrial Competitiveness which emphasized that the nation's ability to compete in global markets depends upon several interlocking elements. One of these relates to improving our nation's competitive strength in science and technology. The enormous development of new products and markets that have resulted from the process of transfer of fundamental scientific findings into commercial use by the application of advanced engineering both within the U.S. and abroad is well known to all of us, and must continue to be fostered in the future.

Dr. SILVER. Thank you, Dr. Mason.

I would now like to introduce Dr. Herbert Simon, who is going to address a number of interesting areas of activity which range from mathematics through the computer sciences into the social sciences.

Dr. Simon.

[A biographical sketch of Dr. Simon follows:]

DR. HERBERT A. SIMON

Born June 15, 1916 in Milwaukee, Wisconsin. A.B., University of Chicago, 1936; Ph.D., 1943. Staff Member, International City Manager's Association, 1938–39. Director, administrative measurement studies, Bureau of Public Administration, University of California, Berkeley, 1939–42. Assistant professor (political science), Illinois Institute of Technology, 1942–45; associate professor, 1945–47. Professor (administration), Carnegie Institute of Technology; head, Department of Industrial Management, 1949–59; associate dean, Graduate School of Industrial Administration, 1957–73; professor (administration and psychology), 1962–66. Richard King Mellon University Professor of Computer Science and Psychology, Carnegie-Mellon University, 1966—. Ford Distinguished Visiting Professor, New York University, 1960. Consultant: Cowles Foundation for Research in Economics, 1947–57; Management Engineering, Economic Cooperation Administration, 1948; The Rand Corporation, 1952–76. Recipient: D.Sc. (hon.); Yale University, 1963; Case Institute of Technology, 1963, Marquette University, 1981; LL.D. (hon.): University of Chicago, 1964; McGill University, 1971; University of Michigan, 1978; University of Pittsburgh, 1979; Fil.D. (hon.): Lund University, 1968; Dr. Economic Science (hon.), Erasmus University of Rotterdam, 1973; Administrator's Award, American College of Hospital Administrators, 1958; Distinguished Scientific Contribution Award, American Psychological Association, 1969; A.M. Turing Award, Association for Computing Machinery, 1975; Distinguished Fellow, American Economic Association, 1976; Nobel Prize in Economics, 1978. Fellow: Econometric Society, American Association for the Advancement of Science, American Academy of Arts and Sciences, American Psychological Association, American Sociological Association, American Economic Association, Experimental Psychology of APA. Member: American Society for Computing Machinery, American Philosophical Society, The Institute of Management Sciences, American Political Science Association, Operation Research Society of America, Psychonomic Society. Associate Editor: Sociometry, 1958–60; Econometrica, 1956–65; American Political Science Review, 1952–56; Behavioral Science, 1963——.

Elected to the National Academy of Sciences, 1967. NAS/NRC activities: Member of Governing Board, NRC, 1980—; Council, NAS, 1978–81; Panel for the Office of Technology Assessment, COSPUP, NAS, 1972—; Committee on the Social Sciences in the NSF (chairman), 1975–76; Committee on the Costs and Benefits of Automobile Emission Control, ESB, CNR, 1973–74; Committee on the Uses of Computers, DMS, 1962–66; Assembly of Behavioral and Social Sciences, 1973–77; Class Membership Committee (Class V), NAS, 1976, 1978, 1982; NRC Division of Behavioral Sciences (liaison), 1963–66; Temporary Nominating Group for Class V (chairman), 1971, 1972, 1973; Executive Committee, ABASS, 1973; U.S. National Committee on the International Union of History and Philosophy of Science (ex-officio), DBS, 1968–70; Executive Committee, DBS, 1966–68, 1968–70 (chairman), 1970–72; Report Review Panel (psychology), Report Review Committee, NAS, 1970–76; Advisory Committee on Government Programs in the Behavioral Sciences, DBS, 1965–68; Committee on Science, Engineering, and Public Policy, NAS, 1967–70; Division of Behavioral Sciences (member-at-large), 1966–68, 1968–70 (chairman), 1970–73.

STATEMENT OF DR. HERBERT A. SIMON, RICHARD KING MELLON UNIVERSITY PROFESSOR OF COMPUTER SCIENCE AND PSYCHOLOGY, DEPARTMENT OF PSYCHOLOGY, CARNEGIE-MELLON UNIVERSITY PITTSBURGH, PA, AND MEMBER, COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY, NATIONAL ACADEMY OF SCIENCES

Dr. SIMON. Mr. Chairman, Members of Congress, the committee's *Outlook* report shows that the scientific community has been very much concerned in recent years with the growing need in science for multidisciplinary research, research that brings together specialists from different disciplines, and with the mechanisms that may make that possible.

Ed Mason referred to one example, the advances in molecular biology and genetic engineering which need to be cross-fertilized, so to speak, with agricultural research. Those two fields need to grow closer together.

The example I would like to talk a little bit about is the modern computer, which has been the subject of about six of our briefings over the past 4 years.

When we talk about the computer, we have to talk about a whole range of things, from the physical architecture, on the one hand; the software, the programs that make it run; artificial intelligence, the attempt to make computers smarter; cognitive science, the application of computers to understanding human thinking, human problem solving, and learning; and, finally, the application of that, in turn, to the educational process, and especially the science and math education.

We need basic research all along that spectrum which brings together the people at the various points along it. For example, we are now getting new, very powerful, miniaturized, faster computers, cheaper computers, and that is being accomplished in part by making our computers parallel devices in which a lot of things can go on at once instead of one thing at a time, for example, in things like visual pattern recognition.

But parallel computers are extremely difficult to program, and the utilization of the new hardware is going to depend on basic advances in computer programming theory that in turn have to depend on progress in basic research on artificial intelligence.

Artificial intelligence, again—that is, trying to make computers "smart"—underlies the next stages of progress in robotics, and it underlies, also, the progress that we have made and are continuing to make in using computers to simulate human thinking and thereby to understand better our own learning processes and to begin to apply that understanding to learning and teaching in the schools.

Here we need collaboration of computer scientists with experts on learning theory and educational psychology and with experienced teachers. We need research, basic research, on the processes that go on in thinking of mathematicians and scientists and the feeding of that into the improvement of science education.

There are important requirements here for help from basic pure mathematics. For example, the mathematical theories we call automata theory and the theory of post-productions were developed originally in an interest in the foundations of mathematics itself.

No one had any idea that they were applicable to anything. And yet today they underline much of the fundamental research in computer science itself—an example, I think, of what has become familiar to all of us in mathematics, of pure mathematics, in unexpected ways contributing to and underlying very important applied development.

We need, also, multidisciplinary approaches to science and technology policy, to the kinds of concerns of this committee, for example, in understanding the problems of energy and the environment.

Fortunately, American scientists in the last decade or decade-and-a-half have opted for collaboration and for unity across the physical and the biological, the engineering and social and behavioral sciences, all of us who are represented at the table here. We have today, I think, most of the institutional framework for that collaboration, such as the NSF. The mission of the NSF has been gradually broadened over the years to include all of those sciences, and within the Academy structure—the National Academies of Science and Engineering and the Institute of Medicine—again, we have a set of institutions that can provide some help and some advice on science policy matters to the Government.

In our judgment, the strong bonds of communication and cooperation throughout the whole community of science are a major asset of American science, both when progress in research calls for multidisciplinary effort and when diverse talents must be combined to address public problems. The preservation and enhancement of that unity should be a major goal of science and technology policy.

Thank you.

[The prepared statement of Dr. Simon follows:]

Prepared Statement of Herbert A. Simon, Member, Committee on Science, Engineering, and Public Policy, National Academy of Sciences, National Academy of Engineering, Institute of Medicine

Mr. Chairman, One of the topics that comes up repeatedly in the Academies' deliberations on science policy is the growing need for multi-disciplinary research, and for the development of effective mechanisms to make it possible[1]. One sucn requirement which has received our attention, is for mechanisms to facilitate application of the spectacular advances in molecular biology and genetic engineering to agricultural research[2].

Another important example, in which the interdisciplinary opportunities are especially broad and multidisciplinary work especially essential, centers on the modern computer[3]. This research domain ranges all the way from the physical architecture of the computer itself, to its software, to the enhancement of its capabilities through the techniques of artificial intelligence, to its use as a basic research tool in cognitive science, to the study of human thinking, and finally the application of the new understanding and insights gained in cognitive science to the improvement of our elementary and secondary education, especially education in science and mathematics.

At each point along this wide spectrum there are vital opportunities for basic research, but some of the most essential research lies in linking together the adjacent specialties, for example, computer architecture with computer software design and with artificial intelligence and cognitive research on pattern recognition and vision[3]. An examination of current trends in computer architecture shows that the rapid miniaturization, speedup and cheapening of computer memories and other components creates possibilities for the design not only of powerful

---

1. The Outlook for Science and Technology 1985. Washington, D.C., National Academy Press, 1985.
2. Research Briefing 1982 (Report of the Research Briefing Panel on Agricultural Research). Washington, D.C., National Academy Press, 1982, pp. 57-80.
3. Research Briefing 1984 (Report of the Research Briefing Panel on Computer Architecture). Washington, D.C., National Academy Press, 1984, pp. 1-16.

"supercomputers" for traditional number-crunching applications, but also of large parallel computers especially adapted to such tasks as visual pattern recognition. For the first time, parallel computers will provide potential computing power allowing at least crude simulation of the human retina.

However, parallel computers have proved exceptionally difficult to program, so that their potential will not be realized without significant advances of our understanding of fundamental programming theory. The programming systems that are required for these computer architectures will have to draw, in turn, upon the field of artificial intelligence, which seeks both to imitate aspects of human intelligence and to cast light on the processes of human thinking and perception[4]. Thus, effective parallel processing will require a fundamental understanding of computational models, algorithms, programming languages, artificial intelligence, optimization techniques, processor-memory-input/output relations and interfaces to existing computing systems and end users.

At the same time, there is rapidly increasing multidisciplinary activity at the other end of the computer spectrum. Computers used in research on artificial intelligence both enable us to produce increasingly sophisticated robotic systems for complex manufacturing tasks[5] and allow us to simulate human processes of thinking, problem solving, and learning[4]. Through the insights we gain from the latter advances, we are beginning to gain a deep understanding of what is required for effective human learning in school situations -- for example, how textbook and classroom materials as well as programs for computer-aided instruction need to be organized and expressed to serve as good teaching and learning tools.

---

4. Research Briefing 1983 (Report of the Research Briefing Panel on Cognitive Science and Artificial Intelligence). Washington, D.C., National Academy Press, 1983, pp. 19-36.
5. Research Briefing 1983 Report of the Research Briefing Panel on Computers in Design and Manufacturing). Washington, D.C., National Academy Press, 1983, pp. 49-64.

Here, the domain now labeled as "cognitive science" brings together computer specialists, researchers on learning theory and educational psychology, and experienced teachers in an exciting joint venture directed toward the advancement of education[6]. Much of the current research in this domain is directed toward education in such subjects as arithmetic, algebra and geometry, physics, chemistry, reasoning, and problem solving, and hence, if pursued vigorously, can soon make an important contribution toward the improvement of science education.

As in so many domains of science, mathematics, including basic mathematics developed without specific applications in view, provides crucial intellectual tools for the researchers trying to think through their complex problems in the computer domain[7]. Two bodies of mathematical theory, known as the automata theory and the theory of production systems, originally developed in research on the foundations of mathematics serve now also as basic means for thinking about the fundamental theory of computers and are under rapid development in the hands of mathematicians and computer sciencetists alike. So, we see another dimension of interdisciplinary work, which depends both on the development of pure mathematics that later provides tools in so many unexpected ways, and the establishment of linkages between mathematics and its domains of application -- in this case, computer science.

The need for multidisciplinary research in many fields, of which I have mentioned only two, and the need for trained scientists who can carry on research that straddles traditional disciplines, is paralleled by the requirements of multidisciplinary approaches when science and technology are applied to

---

6. Research Briefing 1984 (Report of the Research Briefing Panel on Information Technology in Precollege Education). Washington, D.C., National Academy Press, 1984, pp. 17-30.
7. Research Briefing 1982 (Report of the Research Briefing Panel on Mathematics). Washington, D.C., National Academy Press, 1982, pp. 1-18).

issues of public policy. We have learned, sometimes the hard way, that problems of energy and environment, for example, require for their understanding the best present and future knowledge about the physical environment, about biological organisms, and about our own human institutions and human behavior. The sciences -- physical, biological, social, and engineering -- share more than the common name of science. They share also the responsibility for making collaborative contributions to our understanding of public problems and our devising of solutions for them.

It is fortunate, therefore, that the American scientific community has opted for collaboration and unity. Through a series of policy decisions that were taken principally in the late 1960's and early 1970's, the whole community of physical, biological, social, medical, and engineering sciences has been welded together in a set of institutions that make their cooperation easy and natural. One of these institutions is the National Science Foundation, whose range of responsibilities was gradually extended from the core physical and biological sciences to encompass the social and engineering sciences as well. Another is the Academy complex, of which our COSEPUP committee is a part, which similarly can draw upon the whole range of talents in American science to address public problems that are identified by Congress, by the administration, or by the scientific community itself.

In our judgment, the strong bonds of communication and cooperation throughout the whole community of science are a major asset of American science both when progress in research calls for multidisciplinary effort, and when diverse talents must be combined to address public problems. The preservation and enhancement of that unity should be a major goal of science and technology policy.

STATEMENT OF DR. LEON T. SILVER, WILLIAM A. KECK FOUNDATION PROFESSOR OF GEOLOGY, DIVISION OF GEOLOGICAL AND PLANETARY SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY PASADENA, CA; AND CHAIRMAN, COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY, NATIONAL ACADEMY OF SCIENCES

Dr. SILVER. Thank you, Dr. Simon.

I would now like to turn to a discussion of some of the issues which we believe are appropriate for congressional consideration. It is worthwhile at this time to consider the context in which we are appearing before you.

Your committee has undertaken a major review of national science policy at a critical time. Now, in science, the moment is always critical because we are always making new progress. But this is a time of criticality in terms of the national use of its resources, and we all recognize that problem.

The major partners in the research activities—the universities, industry, the Federal and State Governments—are now continually redefining their respective roles as the magnitude and costs of these activities grow. To guide the national science and technology effort, policymakers must ensure that national funding for science is deployed efficiently and productively. Because of the long time constants characteristic of the research enterprise—and I stress those long time constants—decisions on directions and priorities must be evaluated for their long-term consequences as well as for their more immediate returns. Central to any policy decisions must be the recognition that fertile research environments require a continuous flow of the best young talent, excellent facilities, adequate, stable funding, and program flexibility. Given this context, I will turn to some of the identified issues.

One which is very much in the minds of this committee and our community is the international competition in science and technology, and for the foreseeable future, the Nation's economic wellbeing will depend heavily on its ability to compete in the international technological marketplace. Suitable mechanisms for creating and applying new technology are necessary, and these mechanisms must operate in a way which is protective of our creative advantages, but they must not stifle further growth.

In such diverse competitive fields as supercomputers, biochemical engineering, and advanced composites, as we have heard, we find that the United States has achieved basic research leadership. However, the United States is being severely challenged in these areas by other nations. Therefore, Congress must give consideration to policies to increase industry and university cooperation and feedback; to identify the special contributions of the partners and to support them appropriately; to create policies which reduce conflict of interest concerns; second; to make the research conducted in national laboratories more available for applications in high technology industries; again, to provide support for cross-disciplinary research; to encourage development of complementary competences both in the basic sciences and in developmental engineering; and to anticipate the technical personnel needs in critical competitive areas by developing mechanisms to encourage both students and faculty to focus on those areas.

Now, you have already heard some, I think, meaningful comments on the problems of scientific and engineering personnel. Management of our national resources of scientific and engineering personnel presents policymakers with a broad range of issues.

The numbers of quality of successive cadres of young people attracted to technical careers must always be of primary interest to you. Mechanisms to minimize the inevitable mismatches in the supply and demand of young scientists should be considered.

Young peoples' interests—and I am a professor and I have some contact with them—young peoples' interests in careers in science and engineering are responsive to current job opportunities and to current perceptions of personal rewards and support. Thus, the computer sciences currently draw large numbers of students who enter industry, while there simultaneously exist critical shortages amongst faculties in computer sciences.

As we have heard, in other areas, talented young scientists, increasingly discouraged by the intense competition for available research grants and by the lack of permanent positions, are leaving science fields for more remunerative and secure employment. It is a problem facing many of our universities, where anticipated retirements later in this decade, in the 1990's, will create numerous faculty vacancies. We must look ahead and plan now to meet these future requirements.

The training of non-U.S. citizens in increasing proportions calls attention to several trends. About one-fifth of all the doctoral degrees in science and engineering in the United States in 1983 were awarded to students holding temporary visas. In engineering in that year, almost 40 percent of the graduate enrollment was comprised of non-U.S. citizens; 38-eight percent of the doctoral degrees in mathematics and 35 percent in agricultural sciences went to foreign students.

It is in the national interest to provide graduate education to the brightest students, regardless of their origin. These students provide a source of faculty for our engineering schools and personnel for U.S. industry. The concern must be with the reduced proportion of U.S. students taking advanced training, especially in engineering.

The Congress may wish to consider as issues: the need for Federal programs and policies that would help to minimize the mismatches resulting from fluctuations in the supply and demand for young investigators; additional programs to help young researchers begin their careers, including provision of funds for the initial equipment grants, as Dr. Bigeleisen has pointed out; and a review of policies affecting the education of qualified foreign students in American universities, their potential for employment by U.S. industries and universities, and the implications of such policies for U.S. institutions and those of other countries.

Now, Dr. Simon has stressed the importance of cooperative work across disciplines. This is a very important opportunity for us. It is one of the things that demonstrates American flexibility and progress in science, our ability to mix disciplinary efforts in fruitful collaborations. However, academic institutions, conservative, do not easily cross their disciplinary domain boundaries.

Congress may well wish to consider an appropriate role for encouraging multidisciplinary research, extending from the physical and biological sciences to the applied and social sciences. I have to stress that progress may be made by any combination of the scientific disciplines. We cannot always predict where that interdisciplinary progress will be made.

There must be incentives for universities to undertake this intramural multidisciplinary research, as well as incentives to induce collaboration amongst several different classes of institutions. This may include industry, the university, and we are now, in the Academies, currently sponsoring efforts to find the common ground.

Mechanisms for encouraging the fruitful integration of basic biological concepts into modern agricultural research constitute a special case. The Congress has undertaken an improved program of support for agricultural research. We consider this a most valuable contribution. But we still find that there is a delay in the infusion of modern molecular biology concepts into agricultural research, and we think this is a matter very much deserving of congressional attention.

You have heard Dr. Bigeleisen address the problems of instrumentation, and I am sure Congress is very sensitive to this matter. There are some fundamental questions for Congress: How can the presently available funds be utilized most effectively and equitably? What leverage can universities bring to the Federal funding to increase that funding's effectiveness? How can the long-term needs for continuing recapitalization of the research university be placed on a sound financial basis? That is a very fundamental question.

Among the issues that you may wish to consider are: first, enhancing the systems for comprehensive merit review of proposals for large facilities and instrumentation, which would incorporate, as appropriate, judgments from the science community on the quality of science likely to be conducted at a facility to be built with Federal funds; second, anticipating and planning for future needs for large scientific facilities; third, seeking optimal ways to leverage Federal investments in such facilities; fourth, considering criteria to test whether international cost sharing is appropriate for a large facility; fifth, ensuring that support for large new facilities is balanced by adequate support for the cadre of qualified individual investigators who continue to produce the bulk of U.S. science; sixth, reviewing current Federal fiscal and accounting policies with respect to the capital and operating costs of such facilities; and, seventh, evaluating alternate mechanisms for funding instrumentation including capital cost recovery as a part of the operating budget of a Federal grant or contract.

In the current *Outlook*, additional significant issues have been developed for congressional consideration. They include issues in genetic engineering and human biology, as mentioned by Dr. Bloom; research in transportation, scientific communication, technology transfer, and national security, a critical issue at this time; and the global effects on the atmosphere of nuclear explosions. We will be pleased to answer questions on these topics.

In summary, it is my view that with a few exceptions, some of which are notable, the general health of U.S. science and technology is excellent. The national research programs still produce international scientific leadership.

Nevertheless, it must be noted that international science and technology is a dynamic system with many able, well-funded contributors striving for a larger share of the leadership. This was to be expected. The U.S. investment in science and technology no longer overwhelms that of other industrialized nations.

The U.S. role can be sustained if attention continues to be given to the timely provision of the basic resources required for a program of reasonable growth. The American research universities, the climate and career awards for young investigators, superior instrumentation and facilities, cooperation and collaboration amongst the major partners in the national research efforts, and a climate of open inquiry have been the key ingredients for our past successes. They can and should be continually enhanced wherever possible to ensure our future.

In closing, let me emphasize that the scientific and technical communities recognize that these are difficult times. The resources that are available have to be allocated effectively and wisely, both to meet short-term crises and to contribute to long-term national growth. With careful planning, science and technology can contribute significantly to this growth.

[The prepared statement of Dr. Silver follows:]

PREPARED STATEMENT OF LEON T. SILVER, CHAIRMAN, COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY, NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF ENGINEERING, INSTITUTE OF MEDICINE

Mr Chairman, on behalf of the National Academy of Sciences, the National Academy of Engineering and the Institute of Medicine, the Committee on Science, Engineering, and Public Policy (COSEPUP) is pleased to present to the Congress the report: "The Outlook for Science and Technology, 1985". This report is the fourth in a series, called for in the National Science and Technology Policy, Organization, and Priorities Act of 1976.

As was the case for each of the earlier reports, the design of the "Outlook" has evolved in accordance with the formal guidance and informal interactions with the Congress. The content derives from the recent insights and activities of the Academies.

The current report discusses nine areas of science and technology selected for research briefings by the Office of Science and Technology Policy in consultation with COSEPUP. It must be emphasized that selection of these nine topics does not imply that other topics are considered to be of lesser importance to the progress of U.S. science and technology. No priority setting among fields and technology is implied. It then identifies for Congressional consideration a number of policy issues drawn from both the accompanying science discussions and from an examination of the current environment for United States science. The issues represent a selection of concerns and opportunities which seem to us to merit the attention of Congressional policy makers. We have attempted to do this in outline form to attain brevity and maintain interest, as requested by the Congress. In this regard, I might note that the first and second Outlook reports were more than 500 and 700 pages in length, whereas the current report is just 50 pages.

Mr. Chairman, the late Dr. George Low, my distinguished predecessor as chairman of COSEPUP, stressed to this committee in January 1984, the importance we place on a continuing dialogue aimed at improving the Outlook reports. We have received constructive comments from members of your committee, your staff, and other members of Congress. We look forward to your continuing appraisals and suggestions.

Before I turn to several of the issues identified for the Congress in the "The Outlook for Science and Technology, 1985", let me provide some general context for this discussion.

Your committee has undertaken a major review of national science policy at a critical time. The research and development activities of the United States contribute to the health and wealth of the nation, and require an appropriate faction of national resources. The major partners in the research activities, universities, industry, and federal and state governments, are continually defining their respective roles, as the magnitude and costs of those activities grow. Definition of roles and interweaving of responsibilities for implementation of the great diversity of programs, in a constructive alliance, requires constant policy surveillance and effective policy-making mechanisms. To guide the national science and technology effort, policy makers must ensure that national funding for science is deployed efficiently and productively. Because of the long time constants characteristic of the research enterprise, decisions on directions and priorities must be evaluated for their long term consequences as well as for their more immediate returns. Central to any policy decisions must be the recognition that fertile research environments require a continuous flow of the best young talent, excellent facilities, adequate, stable funding, and program flexibility.

Given this context, I will turn to some of the identified issues.

International Competition in Science and Technology

For the foreseeable future the nation's economic well-being will depend heavily on its ability to compete in the international technological marketplace. In addition to supplies of productive capital, an educated labor force and a determined policy stance on international trade, the country must have the ability to create and apply new technology. Suitable mechanisms for protecting our creative advantages without stifling further growth, must be established.

In such diverse competitive fields as supercomputers, biochemical engineering and advanced composites, we find the United States to have an established basic research leadership. With these advantages, however, there still remain in each field many vital activities which need to be accelerated and coordinated to maintain technical and economic leadership. The U.S. is being severely challenged in these areas by other nations. Therefore, policies must be established to:

(1) increase industry and university cooperation and feedback; identify the special contributions of the partners and support them appropriately; create policies which reduce conflict-of-interest concerns;

(2) make the research conducted in national laboratories more available for applications in high technology industries;

(3) provide support for cross-disciplinary research; encourage development of complementary competenences both in basic science and in developmental engineering;

(4) anticipate the technical personnel needs in critical competitive areas by developing mechanisms to encourage both students and faculty to focus on these areas.

Scientific and Engineering Personnel

As noted in the discussion of technological competitiveness, management of our national resources of scientific and engineering personnel presents policy makers with a broad range of issues.

The numbers and quality of successive cadres of young people attracted to technical careers must always be of primary interest. Mechanisms to minimize the inevitable mismatches in the supply and demand of young scientists should be considered.

Young peoples' interest in careers in science and engineering are responsive to current job opportunities and current perceptions of personal rewards and support. Thus, the computer sciences draws large numbers of students who enter industry while there exist critical shortages among faculties in computer sciences.

In other areas, talented young scientists, increasingly discouraged by the intense competition for available research grants and by the lack of permanent positions, are leaving science fields for more remunerative and secure employment. This is a potentially acute problem in biomedical research where fewer clinicians are entering the research programs in universities. It is also a problem facing many of our universities, where anticipated retirements later this decade and in the 1990's will create numerous vacancies. As a nation we must look ahead and plan now to meet these future requirements for qualified faculty.

The training of non-U.S. citizens in increasing proportions calls attention to several trends. About one-fifth of all doctoral degrees in science and engineering in the U.S. in 1983 were awarded to students holding temporary visas. In engineering in that year, about forty percent of the graduate enrollment was comprised on non-U.S. citizens. Thirty-eight percent of the doctoral degrees in mathematics and thirty-five percent in agricultural sciences went to foreign students.

It is in the national interest to provide graduate education to the brightest students, regardless of origin. These students do provide a source of faculty for our engineering schools and personnel for U.S. industry. The concern must be with the reduced proportion of U.S. students taking advanced training, especially in engineering.

The Congress may wish to consider as issues:

(1) the need for federal programs and policies that would help to minimize the mismatches resulting from fluctuations in the supply and demand for young investigators;

(2) additional programs to help young researchers begin their careers, including provision of initial equipment grants;

(3) a review of policies affecting the education of qualified foreign students in American universities, their potential for employment by U.S. industries and universities, and the implications of such policies for U.S. institutions and those of other countries.

Cooperative Work Across Disciplines

My colleagues and I have stressed the great importance and productive power of multi-disciplinary cooperative research. The sophisticated instrumentation of a physicist, constructed by an electrical engineer, may be essential to the successful development of a production process designed by chemists and biologists for some aspect of biotechnology. In this fruitful collaboration each contributor responds from his disciplinary strength to the problems posed.

Industry has invoked this task-force approach with great success. Universities are much slower to activate programs which cross their traditional disciplinary domain boundaries. Mechanisms need to be created which will encourage multidisciplinary efforts without reduction of disciplinary quality.

A special case can be made, for example, for encouraging the application of recent advances in molecular biology and genetic engineering to agricultural research. At present there are not enough individual scientists trained in molecular biology and in agronomy and the basic plant sciences to bring the full power of recombinant DNA and related techniques to agriculture. Wisely, Congress initiated the competitive grants program for agricultural research in 1977. This program is one of several mechanisms which could be used to infuse modern biological technologies into agricultural research. It is currently funded at too low a level to meet the full range of challenging opportunities.

Congress should consider:

(1) an appropriate federal role for encouraging multidisciplinary research;

(2) incentives for universities to undertake intramural multidisciplinary research, as well as collaboration with other institutions;

(3) mechanisms for encouraging the fruitful integration of basic biological concepts into modern agricultural research

Facilities and Instrumentation

Congress has recognized the fundamental issue of the inadequacy of facilities and instrumentation on all scales in our research universities. As it considers appropriate remedial measures it will be confronted simultaneously with the constraints of national budget deficits and the pressures for special political interventions in the appropriations process.

The questions for Congress include:

(1) How can the presently available funds be utilized most effectively and equitably?

(2) What leverage can universities bring to the federal funding to increase their effectiveness?

(3) How can the long-term needs for continuing recapitalization of the research university be placed on a sound financial basis?

These are difficult questions which raise critical issues for the managers of the national scientific and engineering research systems. Among them are:

(1) enhancing systems for comprehensive merit review of proposals for large facilities and instrumentation which include judgements from the science community on the quality of science likely to be conducted at a facility to be built with federal funds;

(2) anticipating and planning for future needs for large scientific facilities;

(3) seeking optimal ways to leverage federal investments in such facilities:

(4) considering criteria to test whether international cost-sharing is appropriate for a large facility;

(5) ensuring that support for large new facilities is balanced by adequate support for the cadre of qualified individual investigators who continue to produce the bulk of U.S. science results;

(6) reviewing current federal fiscal and accounting policies with respect to the capital and operating costs of facilities; and

(7) evaluating alternate mechanisms for funding instrumentation including capital cost recovery as a part of the operating budget of a federal grant or contract.

Other Issues for Congress

In the current "Outlook" an additional number of significant issues have been developed for congressional consideration. Many of them have been raised in the presentations by other members of COSEPUP. They include issues in genetic engineering; human biology; research in transportation; scientific communication, technology transfer and national security; and global atmospheric effects of nuclear explosions. We will be pleased to answer questions on these topics, if asked, in the discussion period.

Summary

It is my view that with a few exceptions, the general health of U.S. science and technology is excellent. The national research programs still provide international scientific leadership. Nevertheless, it must be noted that international science and technology is a dynamic system with many able, well funded contributors striving for a larger share of the leadership. This was to be expected; the U.S. investment in science and technology no longer overwhelms that of other industrialized nations. The U.S. role can be sustained if attention continues to be given to the timely provision of the basic resources required for a program of reasonable growth. The American research universities, the climate and career rewards for young investigators, superior instrumentation and facilities, cooperation and collaboration among the major partners in the national research efforts, and a climate of open inquiry have been key ingredients for our past success. They can and should be continually enhanced wherever possible, to ensure our future.

In closing, let me emphasize that the scientific and technical communities recognize that these are difficult times for the nation. The resources that are available have to be allocated effectively and wisely, both to meet short term crises and to contribute to long term national growth. With careful planning, science and technology can contribute significantly to this growth.

Dr. SILVER. Now, I would like to introduce the Honorable Emilio "Mim" Daddario, well known to the committee, for some remarks about the communication between the science community and the Congress.

Mr. FUQUA. While you are passing the microphone down, just let me say that Mim Daddario was a very distinguished and respected member of this committee and of the U.S. House of Representatives for a number of years. He was one of the first members of this committee and chaired the Subcommittee on Science, Research and Development, which is now chaired by Congressman Walgren, but we have changed "Development" to "Technology." So, it is Science, Research and Technology.

But Mim has had a long career and also, after leaving Congress, served as one of the first directors of the Office of Technology Assessment and as president of the AAAS. We are very glad to have you back, Mim.

[A biographical sketch of Mr. Daddario follows:]

HON. EMILIO QUINCY DADDARIO

Born in Newton Centre, Massachusetts, September 24, 1918. Graduated Wesleyan University, B.A., 1939. Graduated University of Connecticut Law School, LL.B., 1942. World War II Army service in Mediterranean Theatre with the Office of Strategic Services; awarded U.S. Legion of Merit and Italian Medga lia d-Argento. Reactivated with 43rd Division at outbreak of Korean conflict; served as Major with Far East Liaision Group in Korea and Japan.

Mayor of Middletown, Connecticut, 1946–49; Judge of Middletown Municipal Court, 1948–50. Member of the U.S. House of Representatives 1959–71; member of the House Committee on Science and Astronautics; chairman of the Subcommittee on Science, Research and Development; member of the Manned Space Flight Subcommittee; chairman of the Special Subcommittee on Patents and Scientific Inventions; left the U.S. House of Representatives to become the Democratic nominee for Governor of Connecticut in 1970.

Senior partner, Daddario, Slitt, Jacobs, & Sullivan, Attorneys at Law, Hartford, Connecticut, 1952–1973;

Senior vice president, Gulf & Western Precision Engineering Group, Gulf and Western Industries, Inc. 1971-73.

Director, Office of Technology Assessment of the U.S. Congress 1973-77.

Counsel, Wilkes, Artis, Hedrick and Lane, Washington, D.C., 1978-present.

President-Elect, President, and Chairman, American Association for the Advancement of Science, 1976, 1977, 1978.

Member, Council on Foreign Relations

Member, Institute of Medicine Member, Institute of Medicine Executive Committee.

Member, Committee on Science, Engineering, and Public Policy, National Academy of Sciences

Fellow of the American Academy of Arts and Sciences

Member, Joint Budget Committee, National Academy of Sciences

STATEMENT OF HON. EMILIO Q. DADDARIO, ATTORNEY AT LAW, WILKES, ARTIS, HEDRICK AND LANE, WASHINGTON, DC; FORMER MEMBER OF CONGRESS FROM THE STATE OF CONNECTICUT; AND MEMBER, COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY, NATIONAL ACADEMY OF SCIENCES

Mr. DADDARIO. Mr. Chairman, thank you.

I would like my prepared remarks to be placed in the record, if I could.

Mr. FUQUA. We will do that. I noticed several had prepared remarks. We will include them as part of the record.

Mr. DADDARIO. Mr. Chairman and members of the committee, it has been many years since I have been in this room, and I am pleased to be here to participate again in an exchange of ideas between the Science Committee and members of the science community.

There is no better way to handle the whole subject of science policy than the way your committee is doing it at this time, through congressional inquiry. In that process, as demonstrated here today, the National Academy of Sciences, through a series of its presidents over a course of years, from the time when George Kistiakowsky was chairman of the committee, COSEPUP, through presidents Fred Seitz, Phil Handler, and now Frank Press, the Academy has made itself available to give advice to Government which, in the long run, has been helpful in showing the strengths and weaknesses of science and technology in a balanced way.

However, there is a real problem in obtaining truly objective advice from any group, including scientists. They have special interests, vested interests, and much disagreement amongst themselves, and they will also give you a wide variety of stimulating ideas that you will have to choose amongst.

As you do that, I would like to just quote from some advice that I think might be useful, some advice that Dr. Edward Teller gave this same committee some years ago. He said:

> One of your most effective tools is to bring out by appropriate questions and appropriately broad hearings all sides of the question, and then your problem will be which ideas to follow. Do not duck this responsibility. Listen to the advice, and then decide for yourselves.

I would expect, by the time you end these hearings next year, that is not going to be an easy task.

I am pleased to be here again in this room and believe that the report that has been in précis form explained to you is an excellent set of ideas for this committee to examine as it determines science policy for the Government.

[The prepared statement of Mr. Daddario follows:]

PREPARED STATEMENT OF EMILIO Q. DADDARIO, ESQ., MEMBER, COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY, NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF ENGINEERING, INSTITUTE OF MEDICINE

Mr. Chairman, The emphasis you are giving to these science policy hearings by extending them through both sessions of the 99th Congress is a message in itself. You are showing that the Congress is prepared to make a serious and determined effort to maintain our Nation's excellence in science and you are asking the scientific community to help you in the achievement of that goal. Your signal is clear: in order to most effectively promote our country's well being and security the public and the Congress must make a serious and determined effort together to keep informed of the possibilities offered by science, as well as the barriers to its effective use and its risks. That can only be accomplished if there is cooperation between governmental leaders and the scientific and technical community.

. Obtaining truly objective advice from any group, including scientists and technologists, is no easy task. They will have specialized interests, vested interests, and exhibit much disagreement. They will also give you a wide variety of stimulating ideas to choose among. However, when you get to the point where you must make choices, the following advice from Dr. Edward Teller may be helpful:

> "...one of your most effective tools is to bring out, by appropriate questions, in appropriately broad hearings, all sides of the question, and then your problem will be which .. ideas to follow ... Do not duck this responsibility ... Listen to the advice and then decide yourself."

There is no better way to handle this whole subject of science policy than through Congressional inquiry. Through these hearings you will force the disparate parts of the scientific community to give you their view of problems and recommended solutions that are of vital national concern. Your reasoned judgment on what to do is best served by open inquiry within which all areas of science can participate.

Major participants in the development of such an understanding are the scientific learned and professional societies as well as the National Academies of Sciences and Engineering and the Institute of Medicine. As institutions they can play a unique role for they can reach, as in "The Outlook for Science and Technology" before you, a cooperative view of some national scientific needs which require Congressional attention.

It is my impression that members of Congress are now better educated in science and technology than in the past, and I am sure that your staffs and supporting mechanisms are increasingly competent in that regard. Nevertheless, the Congress should not go it alone, and these hearings are a good indication that you intend to combine your increased internal science and technical capability with the opinions and knowledge of the outside scientific and technical community. As an additional step, I would suggest that there be informal contacts between the members of your committee, heads of scientific organizations, and scientific leaders, as these meetings progress, and as you struggle among choices and priorities.

Dr. SILVER. Thank you, Mim.

Mr. Chairman, we are now available for further discussion, as you wish.

PANEL DISCUSSION

Mr. FUQUA. Let me start out with a compliment. I have almost read this "Outlook." [Laughter.]

Subjects have come up during the course of your remarks, and I think you have done a very excellent job.

Dr. SILVER. Thank you.

Mr. FUQUA. It is very concise; it is to the point; it is easy reading; and it is easy to find issues. We established several years ago that it was the quality, not the quantity, that we needed, and we were not getting paid by the pound, and that it would be more concise, more widely distributed and read—and, hopefully, understood—and I think you have done a very excellent job. Let me commend the committee and all of you.

I thought George Low and his group did a very excellent job, which they did. You have likewise, I think, exceeded that in conciseness and really hitting the important points that needed to be touched upon, and I think it is a very valuable tool for members of this committee, but also, I think it will be widely read throughout the Congress and the public and the scientific community.

One of the things that I would like to ask, and those that wish to comment may—we go in session in about half an hour and may be interrupted. It is kind of a two-pronged question. How do you see the priorities of science and the role of science changing in the last 40 years since the Bush report, and where do you see it going in the next 40 years—because that is what we are looking at—and beyond the year 2000?

Dr. SILVER. Let me ask Dr. Press to respond to that.

Dr. PRESS. Mr. Chairman, there is so much forment now taking place in terms of new institutional structures for science, technology, and innovation; there is such unusual progress across so many different fields, that it is very hard to look ahead 40 years.

The boundaries between science and technology today are almost disappearing. One could not have said that 40 years ago. The time between a scientific discovery and commercialization in so many fields today is shortened by a factor of four or five.

Mr. FUQUA. Which is good.

Dr. PRESS. Which is very good.

We just awarded the Nobel Prize this past year—not we but the Swedish Academy—for the discovery of monoclonal antibodies, and within a year or so, it will be a billion-dollar industry led by the United States, incidentally, even though this was discovered by an Englishman.

So these are what characterize science today: unparalleled progress, applications with shorter time scales, spawning of new industries, new ways of doing things. The progress in health care is entirely unpredictable but moving ahead very rapidly.

So there is no question that maybe the principles of doing science today are the same today as they were 40 years ago, and they will be the same in the years ahead, but the state of affairs today is much different, and the future is almost unpredictable but, I think, essentially very, very positive for this country.

Mr. FUQUA. How do you account for, say, a country like England that historically has spent a great deal of money in higher education and research, up until just maybe the last year or so because of budget constraints, and that leads to the problem that they have had a declining economy; and yet a country like Japan that spends very little in these areas but that is one of the dominating countries in the world in trade and technology? How do you explain that?

Dr. PRESS. I would like to turn to some of my colleagues.

Mr. FUQUA. We speak of more and more resources for education and research, and then here is a case of two countries diametrically opposite.

Dr. PRESS. Let me turn to my colleagues, but I would like to introduce that topic in the following way: that, too, is changing very rapidly. The Japanese now recognize that their technological leadership cannot, in the future, continue on the basis of borrowing other countries' science. They are now making special initiatives in key areas of science to feed their technology.

Sad to say, England, I think, is making a very serious mistake in cutting back on its support of science. I agree with my colleague, the president of the Royal Society, who decries these policy decisions of the present government in England at making these reductions; and this at a time when all that I have described before is happening—the boundaries between science and technology eroding, the new industries that are being supported.

So I think these are two countries that are embarking on different paths, one, I think, correctly, and the other, I think, incorrectly.

There are a lot of experts here. Floyd? Lee?

Dr. SILVER. Dr. Bloom.

Dr. BLOOM. We might just look at the English and the Japanese investments in life sciences research. In the life sciences, the English investment has not been effective for the past 50 years. They have not maintained their position in the world with regard to these aspects, and in part it was because they decided that biology was not an interesting field and did not merit increasing support. Except for a very few centers in that country, it has not been possible to maintain their position in the world.

We mentioned the development of monoclonal antibodies in an English laboratory and the development of the industry in the United States. It is true that a discovery does not have to be made in your back yard to exploit it, but had the world of science been slightly different at the moment Professor Milstein described his approach, he might have patented it; and had he patented it, he would have maintained all the rights, and we could not have developed it in the United States.

So it was a partial fluke in the atmosphere of science at that moment that allowed innovative technology to be applied to an area that was somebody else's discovery. It is unlikely to ever happen again, and one cannot wait for the Japanese to pick up on one of our discoveries. We have to be able to implement it ahead of time.

As Dr. Press has noted, the Japanese have made a major commitment toward biological research, and they have come to the United States to try to recruit some of our younger scientists away from us to work in areas where they can provide the equipment right up front. This is a rather major threat to the future of our scientific laboratories.

Dr. SILVER. Dr. Simon.

Dr. SIMON. I would just like to supplement the previous two comments and observe that it, of course, requires more than the basic science research to translate this into economic products and into things of use to the society, and the Japanese showed us that our science could be extremely valuable to a society.

I think we are going to have to focus, and are focusing today, a great deal more on the techniques for bringing our own scientific discoveries into application in our own economy. But I think that once we develop those techniques—or I should not say develop; we obviously have them—once we enhance those techniques, then, as Floyd Bloom points out, the country that does the basic discovery has a real advantage in the ease of transferring that discovery into applications of society.

Word-of-mouth, direct face-to-face personal relations are terribly important in scientific communication and in transferring the skills of science into application.

Dr. SILVER. Thank you, Dr. Simon.

Ed Mason.

Dr. MASON. If I could comment as an engineer, what society sees and realizes is not science so much as it is the applications of science.

You asked a question about the Japanese. The Japanese are now graduating as many engineers as all of the United States, in spite of the difference in population. They are graduating more engineers than all of the degrees in Great Britain. So the Japanese have put a great emphasis on the applications of science, and I think that is one of the reasons they are leading. And, of course, there are factors other than science and technology which affect the entrepreneurial and exploitation aspects—you gentlemen know better than I—taxation, regulatory policies, encouragement by Government and the like.

So we have a whole ball of considerations here that have to be taken into account when one tries to see where should science be done, how much, and, more importantly, how can society, the public in general, benefit from that?

Mr. FUQUA. Mr. Lujan.

Mr. LUJAN. Thank you, Mr. Chairman.

I want to join you in congratulating the group in this publication. I think it is excellent—very readable, very understandable. However, I gather that you have the same problem that we do in this task force, and that is to exactly assign the priority as to where we ought to go in science policy.

It is interesting that you are talking about cell research. I visited two or three laboratories, and I came back convinced that that is going to lead to the eradication of all disease; and if one happens to sneak through, we are going to get it, anyway, because we know just how to do it.

Then you talk to other people about genetic engineering, and they say, "My God, don't do that, you know, you are going to create monsters," and that sort of thing.

So I guess we set priorities here in the Congress by what discipline has a "Godfather" in the committee and in the Congress. We look at the discussions that we have—materials research, for example. Well, we have Godfathers for batteries and steel initiatives and those sorts of things, and so if you've got more votes, you get it—supercomputers, we are all kind of interested in.

Perhaps that all leads to how do we set some strict guidelines, perhaps, that we might say, "Well, it fits into this guideline." The major things being the instrumentation, for example, that all of you could use. We assign one major priority to that; another one, perhaps, to defense, and then subcategories for international cooperation to help them pay for it.

What I look for—and I do not know that it is possible—if we can draw some strict guidelines as to how those priorities can be set, without necessarily mentioning the discipline, what is it that we want to solve? What is it that we want to do? I think that, from my perspective, is what we would like to know from you gentlemen.

Mr. Daddario says we all have our biases. I have been around here long enough to get a little cynical when I look at your nine recommendations and try to match the disciplines to each one of the members of your committee, so I don't know that you make decisions any differently than we do. So can we do this in a strict guideline sort of way?

Dr. SILVER. Well, if we couple some of the points that have been made earlier this morning—the unpredictability of science, the surprising new directions, the relative roles of the major partners, the industrial role, the role of the Federal Government, the role of the State governments—which should not be overlooked in this activity, especially in their support of their universities—and the role of the general science community, you have to create a system which is, if I can pick another phrase, not strict, but one which maintains a constant surveillance over the operations of the general science and technological community and is sensitive to where progress is going.

You do this by creating several mechanisms, most of which are in place. We have the National Science Foundation, the National Science Board, the many panels which advise it on what now seem to be important directions of science, quality, conduct of science, and gives, by virtue of considered review, some sense of priority.

In the administrative problems of segregating funds for instrumentation as opposed to maintaining the operations, this has to be done in a businesslike fashion, which I think is quite clearly in the purview of Congress to establish requirements that the planning be comprehensive both to take care of the ongoing operations and to estimate with some degree of accuracy, always being revised by surveillance, the kinds of instrumentation requirements that are necessary.

When the country has a limit to its resources, then someone has to make priority decisions. As we all recognize, there are advocates in the science community. There always should be advocates in the science community. Congress has the responsibility for making the policy decisions.

When you ask us for priorities, we will give you the best advice we can, but as Mim says, you must consider the source in this thing. We have to help you arrive at what is in the national interest, but I don't think we can lay down any firm, unique guidelines. We have to be responsive to the state of the system.

Dr. **Press**. In the spirit that this is a discussion at this stage, may I try to give you a brief answer to your question?

If you were to concentrate on instrumentation, facilities, and rejuvenation—that is, young people into the system—I think that most scientists would agree that those are the right priorities.

If you were to insist, at a time of tight budgets, that each field define its priorities—what is the most important thing to do, and the second, and the third—I think scientists who have been reluctant to do that in the past would be willing to do that now.

The most difficult question is, how do you compare physics and biology, or astronomy and botany? In the past, there have been almost historic mistakes when we condemned a field prematurely and it turned out to be very, very exciting.

So I would say that if you were to consider core support at a basic level for science in general, and then say are there any special opportunities in this field or that field because of a new discovery or a new breakthrough, new technique for a special push, I think that may be a reasonable way to proceed.

Dr. **Silver**. Jake.

Dr. **Bigeleison**. I would like to caution you about the dangers of getting involved in very strict guidelines and strict planning. We have an example of how poorly that works from the Soviet Union. That is how they plan science in the Soviet Union, and in my view, Soviet science is not very successful. The Chinese tried to copy that system from the Soviet Union and have now realized that is a failure, also.

I just would like to reemphasize what my colleagues have said. The thing to do is to ask the scientists, where is the action? What are the great ideas coming forward? Who are the bright people? And that is the way to go.

Mr. **Lujan**. Thank you.

Mr. **Fuqua**. Mr. Packard.

Mr. **Packard**. Thank you, Mr. Chairman.

There was one comment made, and I cannot recall now in whose testimony, that indicated the possibility that maybe we have reached and explored all the frontiers, or we are at least reaching that point.

In fact, are we almost through with our scientific discovery, and are we only now filling in the voids, or will we, in the next few years, or do you see the possibilities enlarged from normally as discovery and as application of our discoveries takes place, that opens up the field even that much more broadly so that the opportunity for further discovery and for further invention is almost unlimited, even more so? What do you think is the future? Are we reaching the end of the frontier, or is the frontier even greater?

Dr. **Silver**. Well, I think that each of my colleagues would respond in a similar way. In fact, I will offer the microphone. Floyd, I saw you nodding.

Dr. **Bloom**. On the frontier aspects, we are both from California, and development in the United States did not end when the frontier got to California. It just gave us a broader perspective of all the things we had to put together to make the country as it is today. And the country is still developing.

So science is the same way, sir. The more we discover, the more questions we can ask about things that we can learn, about things that we can apply, about things that we can learn and things that we can apply. Science will continue to create knowledge, and with that knowledge will come insight, and with that insight will be applications and opportunities for new jobs, new industries, new ways to prevent disease and promote health.

I think that any of us here would say we are far from the end of the goals of the frontier that science has opened.

Dr. **Silver**. Dr. Mason.

Dr. **Mason**. As somebody from industry, I don't think we are approaching the end of the frontier at all. Many companies now are, themselves, beginning to do a good deal more fundamental or basic research because we find that the more we understand about basic science, the more applications we will have. Each new discovery, if you will, spawns others. The development of optical fibers and now solid state semiconductors which can lase are opening up a whole new field.

People are beginning to talk about computing with photons. When I took physics, a photon was something that had no mass, hardly any properties at all, and now we are finding that people are referring to them the way people used to refer to electrons.

I see continuing wave after wave of—kind of like an artichoke; you keep peeling off and you get closer and closer to the heart, but it keeps growing—as well as the exploitation of the science that we are discovering. I find it is an extremely fruitful field from an industrial point of view, and certainly it is an intellectual field.

Mr. **Packard**. Let me ask one more question, if I may, Mr. Chairman.

In your testimony, Dr. Silver, you alluded to the disparity between American and foreign students as they make scientific pursuits. And I noted a bit of concern in your testimony about that particular area. I, frankly, am not certain what we ought to do in this country in terms of a national policy in regard to the training and the education of foreign students versus our own. We have historically captured many of the brightest foreign minds and brought them here, and they have become extremely valuable in terms of space, biomedicine, and a multitude of other scientific pursuits. Therefore, I am wondering if it is a wise policy. You bring up the question as to whether the Congress may wish to consider some kind of disincentive for foreign students to be trained in our own schools. Or, on the contrary, should we invite them and then encourage them to remain and work in our applied science areas?

Then, leading back to the chairman's first question, or one of his questions, which I was also interested in, are we doing an adequate job in this country of actually applying the science and converting our scientific research into an economic benefit to the country? Because, frankly, if we don't, if we allow other countries to move ahead—and Japan has been a good example of that, and England has been a good negative example of that—where we fail to apply into a productive product or a productive economic benefit for the country the research that is being done, that probably is what will invite more and more scientists and young minds to come into the scientific field.

Perhaps that does not cross, or does not need to cross, international boundary lines. I would like some further comment from members of the panel on this whole question of foreign students and also the application as it has been alluded to by the chairman.

Dr. **Silver**. Well, Mr. Packard, you have phrased, I think, several of the considerations very well. Much of U.S. science has drawn on the brilliant minds of the world and has brought many of them to this country and has grown because of the benefits of that. I don't think we see any change in the desirability of having, in American universities, the finest minds of the world that we can draw and attract. We will grow in our science with the presence of that participation in our scientific community.

We are not suggesting that there are problems with having too many foreign students. The concern is that in many key disciplinary areas—engineering, for example—that young Americans do not perceive the attractiveness of careers in this area, and that is where our concerns are. I think that is a matter which Congress should address its attention to.

Now, I know my colleagues also share views on this. Is there anybody who would like to comment?

Dr. **Bigeleison**. There is another issue that the Congress needs to consider with respect to foreign scientists. Educating foreign nationals here in the United States is a very important positive aspect for building bridges toward the future in our foreign policy and also exposing the American population to other cultures.

Now, the problem which I believe you need to consider is the question of what happens to the foreign scientist who comes here and then stays here. That person makes a contribution to our education and our economy. But when that person does not return, that person does not help improve the country he or she came from, and we are then adding to a brain drain from that country.

It acts in a negative way for our goal to try to improve the underdeveloped countries. So there clearly has to be some kind of optimization of this problem, and you in the Congress are the ones who have the greatest insight into this problem.

Mr. **Packard**. Thank you, Mr. Chairman.

Mr. **Fuqua**. Mr. Walgren.

Mr. **Walgren**. Thank you, Mr. Chairman.

I don't have a very deep background in this, and knowing that Mr. Daddario had preceded me in this subcommittee as chairman makes me question whether history is inclined in the right direction. But as I understand it, it is said that most of the studies of science and its accomplishments with respect to the effect that we put into it are relatively not quantitative or based in retrospective analysis but rather are future oriented and focused on holding out the great possibilities in each area.

Somebody said that most countries have better statistics on poultry production than on the activities of their scientists and engineers. Is this a wrong criticism of our view of science at that point? Should we have a more comprehensive evaluation of the results of the investment that we have made rather than being satisfied with individual stories of great breakthroughs that in fact have a magical quality to them but are not focused on the broad impact of the dollars that we have invested?

Dr. SILVER. Dr. Bloom.

Dr. BLOOM. One way to look at that question might be a couple of analogies from everyday life. We don't really want to count the number of papers as though every paper was an equivalent unit, so that we count them like eggs as they get laid on the farm.

What really counts is the breakthrough. There are an awful lot of feelers that have to exist before there can be breakthroughs, and what counts is the momentum and the attraction that the new knowledge brings to the field by drawing other scientists from misdirected efforts or from not correctly directed efforts to a line where they can all contribute to what appears to be a verifiable fact.

So we don't count our papers as equivalent to knowledge. What we count as knowledge is the breakthrough, is the ability to put together facts that previously were unconnected in a way that now makes an hypothesis that can be experimentally tested, shown to be—at least for the moment—true, and allow those people to put that information together and draw new experiments that will carry it forward.

It is very much like a ground battle in a war. There are an awful lot of individual skirmishes, and finally there is a breakthrough that moves the frontier one direction or another, and that is, in a way, the way scientists battle nature trying to make a breakthrough, trying to get a peek at some piece of information that was not available before. Then we rush to that spot and pour all the effort we can muster into that spot.

And so I think my answer to your question, sir, would be no, we don't want that kind of quantitative or retrospective, unless you are unhappy with where we have gotten. What we are really worried about is where we could be rather than how long it has taken us to get to where we are.

Dr. SIMON. May I add a word on that?

Dr. SILVER. Dr. Simon.

Dr. SIMON. I guess I make some distinction between the quantitative and the retrospective. I agree entirely with Floyd on the quantity. There are very few things you can measure quantitatively here that are worth measuring in that sense, and yet we do have a considerable sense of importance in a more qualitative way.

We all knew when we got to the Moon. That was a goal we set, and we reached that goal, and we thought it was an important goal, "we" now being the whole American people.

I think we also have some real measures of whether we are making progress in medicine as a result of basic biological research. And within science, too, when the breakthrough comes, or if it doesn't come, the young and able people who are going into science do sense where the action is. They are in many ways part of the rudder of this whole scheme because they tend to push and to go in the directions where there is movement, and maybe sometimes even the older ones among us are smart enough to do that, but mostly the young.

So I think that kind of measurement does go on. Now, our chairman complimented us today on reducing a 500-page book to 50 pages, and I think here, too, we really have to be concerned with quality and with relating what we do to social goals and societal goals and intellectual goals in that qualitative way, and it is pretty hard to count.

Mr. **Walgren**. If I could ask one more question, Mr. Chairman, this sort of strikes me as kind of two teams here, we on this side and you on the other, and what I hear is the ball sort of moving from your side to our side in terms of this decisionmaking process, or at least some of the phrases that we have used here this morning.

For example, Chairman Silver said that Congress must make the policy. You can ask us for priorities, but you have to consider the source. Someone indicated that scientists have special interests, and Dr. Press emphasized that it is very hard to decide among the fields, that priorities are easier to set within a field, and we can tell you where the exciting things are, but you should not abandon any field, and you certainly make choice among fields at your risk.

Dr. Teller says that you are supposed to decide for yourselves after reviewing as comprehensively as you can. And yet we are at a time where there has been some specific criticism of our deciding for ourselves from time to time. This has been particularly with respect to the Congress targeting some facilities, as you know.

In your report, you indicate that you need to enhance systems for comprehensive merit review of proposals for large facilities which would include the judgments from the science community on the quality of science likely to be conducted at a certain facility.

I guess I have two questions. One is focused directly on this concept of the quality of science likely to be conducted at a facility to be built with Federal funds. I am sure things cut both ways, that in a sense it is important who is looking at whom and who is talking to whom. At the same time, we all know that there was no science being done at Kitt Peak until the Kitt Peak Observatory was built, and then a lot of science was done there.

As I understand it, when there are unique facilities, they are going to attract the people, and in fact some of our people are going to France to work on certain supercomputers, and I gather the work on those supercomputers will be of very high quality.

What is the foundation for the idea that there will be a real difference in the quality of the science done based on the location of a facility? And then the second question is, whose decision is this? Is this a decision for the scientists to make, or is this a decision for some other body, all with our drawbacks, but nonetheless, some other mechanism? Can it be left to the scientists?

We talked one time about perhaps some liberal arts approach to deciding some of the larger commitments of funds, which would at least give us confidence that simply the special interests of a relatively narrow group of scientists, or simply the counting of the

Godfathers or the votes on a certain committee, would not be the determining factor.

If I could ask for responses to those two areas, I would be appreciative of your contributions.

Dr. PRESS. Let me start with the first question. I think we on this side of the room recognize that when it comes to large facilities, that when it comes to bricks and mortar, there are many factors that go into a decision. There are political factors, regional factors, economic development factors, all of that.

What frightens us is the bypassing of evaluations, even using the congressional process of evaluations. We are very concerned about decisionmaking by amendment on the floor of Congress, bypassing congressional committees. We are concerned when the congressional committees themselves don't have an evaluation about the scientific importance of a particular facility, the quality of work that is likely to go on there.

So, what we propose is a process that is a traditional process that Congress itself has used over the years; namely, get the best scientific advice about whether this facility will enhance that scientific field. What is the quality of work that is potentially going to take place there? And then you have to factor into that the other considerations by these kinds of hearings, of geography, of politics, of regional development.

But it is the absence of that series of evaluations and the decisions that are made on the floor of Congress, bypassing all of these things, that really make us worry about large-scale diversion of research moneys—most of these facilities do not carry new moneys—without a process of judgment that has been traditional and that is responsible for the great strength of American science.

Mr. FUQUA. If the gentleman would yield, there are cases—and two come to mind—that I think have been initiatives of this committee. One was in robotics several years ago when you were at the White House. That did not come from the administration; it did not come from NSF; it was something that we had determined that was very important.

Another was in the field of supercomputers that was an initiative that we started, and finally, NSF agreed that that was a good program. They put very modest sums, and under the leadership of Mr. Walgren, that was increased significantly.

So, there have been issues and areas that have been congressional initiatives. We did one in math and science and engineering education that was totally a congressional initiative. It was finally signed into law, but it was an initiative that started in the Congress, and a lot of them in this committee. So, there have been programs that had political origins.

It is like the man in the Florida Legislature one time who said, "There is too much politics in politics." It is difficult to sanitize politics and to remove it, just as Mr. Lujan was talking about the Godfathers of various disciplines and laboratories and so forth that support those very strongly, or different fields that they feel strongly in.

Dr. PRESS. But, you see, Mr. Chairman, that is precisely the process we endorse. The way you did it was the right way. It was successful. You were very prescient. You picked two fields that turned out to be of great significance to the country. You took a great deal of testimony, you listened to a lot of witnesses, and you made your decision about these fields.

That is the process that we endorse. That is a deliberative process. You did not do it late at night on the floor of Congress by an amendment. You went through a process that everyone respects.

Mr. LUJAN. Mr. Chairman, if I may, I don't want to just brand anybody as a Godfather. I have been accused of doing that, myself, sometimes, but we are all, I think, Godfathers in this. [Laughter.]

Mr. FUQUA. Make sure that is not an ethnic slur, either. [Laughter.]

Mr. WALGREN. I would like the record to show that my wife is Italian. [Laughter.]

Mr. FUQUA. Mr. Stallings.

Mr. STALLINGS. I appreciate the testimony. It has been very educational for me.

I have a concern, I have a sense that I really cannot put my finger to, but perhaps you could respond to it. I see sort of a two-track development in science in the last few years. On the one hand, business and industry have picked up a big part of it. It is becoming very profitable, and many of the new technological developments have come as a result of industry research, and there are great profits attached to it.

On the other hand, I see the State universities suffering because of various reasons including the higher costs of instruments, the falling tax dollars on the State levels.

Now, it seems to me that the science that is generated by the business community is very end-oriented; in other words, they are looking for things that will maximize profits. The science that is being done in the universities, where in the past it has been, I guess, a generic nature, is going in all directions and just inquiry for inquiry's sake.

What does the future hold? Someone criticized the Soviet system because it also has sort of an end view. Are we going in that direction because of economic considerations rather than because of political/philosophical considerations? In other words, will our science in the next few years be directed by the business community because of the dollars attached to it, and therefore the science that is coming from the universities, from this body, will hurt because of the lack of dollars and the fact that we cannot pick up the cost on the scale that ought to be done?

Dr. SILVER. Dr. Bloom?

Dr. BLOOM. That is a two-track question, so you get a two-track answer.

The profitability issue of the biological corporate world I think is greatly overestimated. Most of the companies now show such little profit that many of the early companies are going under. A lot of the profits come in the stock market side of selling the stock in the company and the people who invested in it, but the people who actually work there have not been able to sell very much yet on a big enough scale to really make it profitable, although we all believe that that ultimately will be the case. Many of the things we know how to make are not yet useful for anything, and so the needs are largely in the research world.

As far as the fundamental ideas and profitability directing the scientific development, I think that is an important concern that we will all want to monitor, but there is no sign that there is a letup in the fundamental pursuit of knowledge for knowledge's sake from which people with the right bent of mind can see a possibility of applying some of that new knowledge to the development of a product or a new application.

Both tracks are important tracks to maintain the motivation to do important research and to do high-quality research, and that issue is one of concern and one that we think currently is working very well, but it is appropriate for you to keep your eye on that series of bilateral developments that are moving in parallel, not in conflict to each other.

Dr. Silver. Dr. Press.

Dr. Press. At the present time, something like 90 percent of the university research is supported by the Federal Government, and perhaps 10 percent, by industry. I am not counting professorial salaries in State universities. Of course, the States make a very significant contribution. But the predominant source of support for the instrumentation, the assistantships, the research time, comes from the Federal Government.

I don't think that is likely to change. I don't think industry is likely to capture a larger fraction, and I don't think the Federal Government, even in times of tight budgets, is likely to reduce in a major way its support of the research that goes on in our universities.

So, I am not worried about a two-tier system or a two-culture system. In fact, I am very optimistic because that 10 percent of industrial linkage with universities is a very important 10 percent, because that is what enables our economy to recapture, for benefit of the whole country, the large Federal investment.

The new linkages between industry and universities, with the Federal Government behind as the patron of basic science, is a unique American invention that I think will do us well in the years ahead. I think it is the envy of all of the other countries in the world.

Mr. Stallings. Let me address this to you, Dr. Bloom. Recent newscasts noted promising developments in controlling of MS [multiple sclerosis] using—I may not be saying this right—monoclonal antibodies. Now, this disease is a great crippler of young adults, since it often strikes people in the prime of their life. This seems to me to be a tragedy. Are we doing enough in researching in that area? Does this development really look promising? Should we be doing more?

Dr. Bloom. As far as I know, sir, the breakthrough that was suggested by that headline article is not one with wide application, and the way in which the monoclonal antibody can help absorb things that may leak into the bloodstream or into the spaces around the cells of the nervous system is probably not going to answer the question as to what causes multiple sclerosis and how one might prevent multiple sclerosis from felling other young adults in their prime of life.

There is a great deal of research being done on multiple sclerosis and on many of the other degenerative disorders of the nervous system, but we don't yet have an insight into the nature of the fundamental process of the disease. There may be many causes of it.

If we substituted the word "multiple sclerosis" for the word "cancer," we could experience many of the same discussions that have gone on in the field. There are high likelihoods that something like a virus at certain times might change the way the immune cells of the body look at the cells of the nervous system and now start to react against them and to cause them to break down.

The disease has a lot of ups and downs spontaneously, and it is very difficult to know whether a new treatment really is a treatment or whether you just happen to start it during a time when that person's disease was going to remit by itself. It is a very difficult area to know with certainty that a new treatment is going to be an effective one, and I would say that none of us would be happy until we not only have a treatment for those that have it but a way to prevent new ones from taking it.

Mr. STALLINGS. Thank you.

Mr. FUQUA. Mr. Brown.

Mr. BROWN. It is a pleasure to see all of you distinguished gentlemen here this morning, most of whom I have seen many times before.

In the number of years I have been in Congress, I have always taken some pride in being an ardent supporter of a strong Federal role in research and development in support of basic science and so on, and yet I always have in the back of my mind this feeling, am I doing the right thing?

I have shared that with many of you from time to time, and it is reflected in this kind of question, which is rhetorical, and I will amplify on it a little bit. But has the $100 billion we have spent on basic research in the last generation, or the approximately $1 trillion we have spent on R&D, made us any wiser, any happier, any more secure?

I point to some current investments. You, Dr. Silver, pointed out the importance of multidisciplinary research applied to agriculture in the field of plant genetics. We have a program in that area, reached after considerable struggle, and still insecure, on which we will spend $250 million over the next 5 years.

We have another program on which we will spend $25 billion over the next 5 years to investigate the possibility of putting a defense system into space against ballistic missiles.

There are two orders of magnitude difference, and there is a question about whether the value of the two programs is really two orders of magnitude different. We have devoted a good fraction of that money that I referred to earlier to fundamental biological research which has extended human life.

The figures show, however, that most of that money goes into extending the lives of people between the ages of 80 and 90, and that we are decreasing our expenditures on preserving the lives of infants, and that our rate of infant mortality is considerably worse than many other advanced countries.

Now, how can this expenditure make us wiser? Why is it you come to us and say it is a good thing to support this thing, all this basic research; now you take it and do anything you want, and we end making a mess out of it? [Laughter.]

Can you speculate on that for me?

Dr. PRESS. I will start, but that is the kind of question where you are going to get eight answers. How many are we here? Seven answers.

Mr. BROWN. Well, we are looking for the wisdom that I asked about.

Dr. PRESS. I am not sure you will get it, but we will try.

Are we wiser? Are we happier? Are we more secure? From one point of view, the specter of nuclear war, I suppose that is a concern of all of us. But I would say that, basically, it makes us optimistic for the future, aside from that terrible issue of nuclear war.

One can look at this in a number of ways, and Dr. Bloom can comment better on what I am about to say. Look at the statistics for coronary disease over the past 20 to 25 years. They are down, mortality due to it. Such things as childhood leukemia; essentially a solvable problem today. Some forms of cancer and their treatment; the worldwide eradication of smallpox—a great scientific advance.

The fact that India is self-sufficient in food, a country that was racked with famines in our lifetime, or China, for that matter, based upon a new agriculture, new forms of crop strains that were developed, that whole green revolution.

The fact that 60 percent of the growth of our GNP is due to new knowledge, mostly due to science; that is translated into economic security, affluence of our own people.

If you want to go back a little further, the average lifespan was 35 years of age 300 or 400 years ago, and, essentially due to science, medicine and health, look at what we enjoy today. Science has brought us a great many problems, but it has brought us a view of a future, I believe, that is more secure economically, health-wise, than we could have expected otherwise.

And so your question is a very profound one, a very philosophical one, and you must never let us forget these issues. But basically, I come down on the side of optimism, and I think if you were to take a poll of the American people, faced with all of the problems and all of the insecurities that they have to look forward to, they would basically come out and say that science is the best way to tackle some of these problems, the major ones.

Who else would like to try?

Dr. SILVER. Dr. Bigeleisen?

Dr. BIGELEISON. Congressman Brown, Mr. Chairman, and other Members of the Congress, in addition to the remarks that Dr. Press has made about the great benefits that science has brought to extending human life and the quality of human life, I would like to read into the record at this point a few extemporaneous remarks that were made by our distinguished colleague, I.I. Rabi, a little over a week ago, when he was awarded the Public Welfare Medal by Dr. Press. These were extemporaneous remarks by I.I. Rabi, whom all of you know.

> Fellow members of the National Academy of Sciences, ladies and gentlemen, this is indeed a deeply moving experience for me to receive this medal, not as a result of

any particular discovery or deed, but a medal for what science has accomplished and, more, what science can do.

This medal points up something for which I have tried to promulgate for all these years, a certain value of science, a spiritual value of science—science, which, in my belief, is the highest achievement of the human race, something which has come to be and is very far from fruition in these last centuries.

Science presents us with a problem and a challenge. Can we, as mortal human beings, born naked in the world, acquire these tremendous powers both over nature and ourselves and still survive, or will our future, like that of Prometheus, be bound to a rock because we have entrusted to a humanity which is ill prepared for it these tremendous powers over nature, over destruction, and, even more, over self-destruction? That is our problem.

I am so pleased with this award which enables me to say these words to our fellow scientists.

That is a quotation from I.I. Rabi. It was a very moving experience for everyone in that auditorium that night.

Dr. SILVER. Dr. Mason.

Dr. MASON. There have been several questions asked about what have we got out of our support of basic research in the past, and just so you will not think we are all thinking the same way, this is a question which is asked very frequently in industry.

A company not only spends money but it has to account for it. I am constantly asked, "What have you done for me lately?" and the only part I don't like about it is the "lately," because the time span between when basic research is done and when the results come forward are not apt to be in the short term.

My company, other companies—and Frank Press has just made general reference to this—are able to look back and see what has developed out of research and development in the company over a long period of time and then equate that to the expenditures over, again, long periods of time. Maybe you spent money 10 years ago and you see the result now. So in making studies of evaluation, one has to be very careful to recognize that. But Frank mentioned the increase in the lifespan of people. One looks back at the NASA research and our advances in air transportation. We look back at the work done on solid state back in the early fifties, and transistors, and look at the electronics industry now. One has to take a very long view.

So I think it is possible, if you will, to show a great return on basic research as long as one does the study properly.

Now, as to the other question, that is not a very good guide as to where money should be spent in the future. The past is past and the future is ahead of us, and I think this question of advice from the scientific community is extremely important.

I feel my role in my company is to make the opportunities—as, also, the threats to my company—available and known to my management, and they, like you gentlemen, have to decide where should money be spent, because they have to take in other factors much as you do.

Congressman Brown, you asked the question basically between expenditures for defense and expenditures for support of life. We can advise you on what is possible to extend life and to make life more bountiful. We can also advise you on what is possible in defense. But when it comes to the decisions, we are just citizens. We don't have any special expertise in where these things should be applied, and I think that is why the question. We are not trying to throw the ball back at you. We are trying to share the bouncing of the ball, in a basketball sense, with you.

Scientists can tell you what we probably can do in the future. Society has to decide what is to be done in the future. I also believe that we can show there has been great benefit to humanity.

Professor Rabi's remarks—one of the things that makes me proud of being an engineer—look back in the Middle Ages. Who could enjoy the quality of life but the people who lived in the castle on the hill? Everybody else was working 16 hours a day trying to grub out a living.

With the development of the industrial revolution, which was mainly mechanical, in the 18th century, and later on in the chemical revolution, toward the end of the 19th century, man now has much more leisure time on his hands, so he can enjoy life much better, and I think that is where science has brought us.

I think it is possible to make a quantitative evaluation. I am not sure it is necessary because there are these qualitative evaluations that can be made.

Thank you.

Dr. **Silver**. Dr. Bloom.

Dr. **Bloom**. Congressman Brown, I will just give you another of the eight answers as to why it is yes for all three of your questions.

Wisdom is one of those qualities that is only slightly harder to conceptualize than memory, but memory tends to be rather short.

One of the reasons that things like Legionnaire's disease and AIDS scare us so much is that they are epidemics, and we have forgotten how frequently there used to be epidemics. But it is with the development of medical technology and health technology that, for the most part, we can control such epidemics once we can understand the causative agent.

You mentioned the fact that we can extend the life of the elderly, but we are not doing very much for the life of the young. But when we look into those records, what we see is that, in the United States, babies tend to be born much earlier, and particularly from mothers who have not had adequate prenatal care; that we deliver babies too early, and we apply an awful lot of high-cost technology to them, when what we really need is to extend the pregnancy in the womb so that those babies come to the birth process at the appropriate time more fully developed.

The infectious disease example and the maternal example are fairly well known, but it is not often realized that in the field of very soft diseases like mental illnesses that the application of medical research has made a tremendous economic impact.

If we just take a simple drug like lithium salts and apply it to the number of people with manic depressive illness who are able to conduct themselves in appropriate jobs and fully lead normal lives, it amounts to billions and billions of dollars that would have been spent on costs of keeping them in hospitals, but they now contribute to the growing marketplace in the economy.

So I think there are innumerable examples we might offer you, sir, but I think the investment has undoubtedly been a wise one.

Dr. **Silver**. Mr. Brown, you asked for wisdom from our side, and as Dr. Press said, you may not get it.

I want to deal with the problem that started about a few million years ago. It was an unexpected infestation and proliferation of a particular species on the planet Earth, and that proliferation has inevitably led to changes in the environment both for that species and all the other species in the system. And with those changes is a whole host of problems.

We try to meet those problems. We find the wisest people we can and try to elect them to Congress, and then some of us go into science and try to do the best we can. [Laughter.]

Those problems are going to be with us. Now, the question is, is science doing its share to meet those problems? And I don't think we are in a position to give you hard answers to the question.

Mr. BROWN. That is the fundamental question that I was getting at.

Mr. DADDARIO. George, I have stayed out of this dialogue because I am not a scientist, and I am being educated as same as the committee is. But you and I have sat side by side on this committee for many years, and I have always enjoyed the philosophical content of your questions, and they have added so immeasurably to the work of this committee.

But I do think that your question implies that science and technology are responsible for our being less happy, not so wise, and not so secure, when in the long run, science policy is only a part of everything that our society is composed of. It is just a tool in the achievement of our other national goals and objectives, and certainly it needs to be worked into the whole framework of everything government does.

To answer your question, we would have to answer policy questions which go into every aspect of American life, science being just one part of it, and I do think that the answers given here show that science adds immeasurably to the opportunity to achieve those objectives.

Dr. SIMON. You asked for eight pieces of wisdom, such as they are, so I am going to add mine, and it is really a footnote to some of the others.

We have a choice, but it is a kind of drastic choice. We can go back to those castles. Probably some of us would be peasants. We can go back to that preindustrial revolution and then we don't need science. But if we want the benefits of the industrial revolution, then we badly need science, not only for wisdom but to survive in this world, as Lee Silver suggested, because our present world is a world with problems as well as with a higher quality of life than any civilization has had before.

And it is only knowledge that is going to solve those problems. It is knowledge of where we can get energy without destroying the environment, and knowledge about ourselves, knowledge about human beings and how they have to conduct themselves and how they have to build their societies in order that they can solve these very difficult problems.

Now, that is going to be hard work and lots of deep and maybe some shallow thinking, but it seems to me that the alternative is very clear. The alternative is that we not merely stand still and take the science we have today, because we cannot have that science without having solutions of some of the problems that in fact are created by our present way of living. We have either got to go forward or we have got to go back, and I find it difficult to perceive just how a society makes that latter decision.

Mr. **BROWN**. Thank you very much, gentlemen, for that response. It has not alleviated my concerns, but it has helped to stimulate my thinking a little bit.

Thank you, Mr. Chairman.

Mr. **FUQUA**. Mr. MacKay.

Mr. **MACKAY**. Thank you, Mr. Chairman.

I worry that our horizons are not long enough, and I see your report saying to us that we run great risks if we bypass the collegial mechanisms that are set up within science to do peer review. I strongly share your view of that, that it is a mistake for us, based on political muscles or other things, to start bypassing that and, in effect, setting our scientific priorities politically.

It seems to me it is fair for us to say back to you, what self-evaluation mechanism do you have to be sure that if we are going to stay out of it, that you are giving us wise advice? I am now using the word "wise" in a sense that is somewhat like Mr. Brown.

It seems to me that your evaluation, your 5-year plan, is fine for 5 years. But what is 5 years? And who is it that is looking at 100 years? And who, if not you, is going to do that? I want to ask, should there not be—I realize we said to look 5 years ahead, and it is quite an indictment for us to think that 5 years is long term. Frankly, 5 years is long term.

But I worry about the implications of our going in the directions we are going and not looking at the off-balance-sheet liabilities that we are building up. Are our soils being sterilized along with our lakes by acid rain?

Now, we have a difficult time dealing with that, not because people doubt acid rain but because we don't want to face the politics of whose fault it is. But should you not be funding research or warning us that we need to fund research in that area?

Dr. Simon, I saw reference to an article that you had written in *Science*, where you talked about the secondary implications of the nuclear winter. I would ask, are we fouling our nest? And should we not, in addition to allowing economic considerations and our own—one way to say it would be humanitarian considerations; another would say we have got a fear-based research policy—should we not have something in there that deals with more long-term considerations? And should you not be allocating money for that and up here defending those decisions? That is my question. Don't we have a major gap in the way we are doing it now?

Dr. **SILVER**. Dr. Press.

Dr. **PRESS**. Let me start and then also turn to my colleagues.

The issue you have raised has been of great concern to me. In Washington, it is very hard to get attention about an issue that is more than 5 years into the future. Not very many organizations here that have the capacity or the interest to address those questions.

Nevertheless, a prudent society is one that does address them, and we are going to try to devote an increasing amount of our resources to this very question.

As you know, most of our funds come from the Federal Government, not as a line item in the budget but as 200 separate projects, and we cannot find agencies to support this kind of issue, so we are trying to raise private monies to do precisely that.

Our major study on acid rain was not supported by the Government; it was supported by private sources. Our concern about carbon dioxide and climatic change is a very deep concern for the future. Our world food studies—where, how will we feed the world in the year 2000 or beyond with this tremendous acceleration in population growth? The whole issue of nuclear winter and its strategic implications. These are on our agenda, but they are only the tip of the iceberg of a host of issues that we have to be concerned about as a prudent society.

Mr. MacKay. The ethical dimension of science.

Dr. Press. Exactly. That is also on our agenda. So we hope, and we hope others will also raise the kinds of questions you do. But it is not easy to gather either resources or an interested audience, and it is not easy to do it well.

There have been a number of futurism-oriented studies that have been flawed, but nevertheless, if science were to change our society, as you have heard from this table, we want to know all of the implications, and I think you have raised a very important issue.

Mr. MacKay. Frankly, I would hope, Mr. Chairman, that our being involved in this whole look at science policy would indicate a higher sensitivity on the part of the political system, that we have got a responsibility to develop a longer horizon and not to treat the person who tries to warn us the way the prophets were treated in the Old Testament, and right now, I think some of what you say, very politely, is that he who plays the role of a prophet gets stoned; he who does not get rewarded.

Dr. Press. To a significant extent, that is right.

Mr. MacKay. Thank you, Mr. Chairman.

Mr. Fuqua. Thank you very much. It has been most beneficial, and I want to thank you, Dr. Press, for being here, and Dr. Silver, you and your committee, for a very fine job. It has been most stimulating.

We do have a hearing scheduled for the politics of science policy late in June, and it should be a very stimulating discussion about that.

But I think it has been very helpful, and thank you all very much for being here and sharing with us your wisdom.

The task force stands in recess.

[Whereupon, at 11:43 a.m., the task force recessed, to reconvene on Tuesday, May 21, 1985 at 10 a.m.]

[illegible]

APPENDIX 1

Questions and Answers for the Record

Dr. Jacob Bigeleisen

1. Your report begins with two highly important and crucial questions: "Of current and potential endeavors in science and engineering, what criteria should be used in selecting projects to be funded? What determines appropriate funding levels?" Can you offer any suggestions about what the answers to those questions are, or how we can go about developing the answers?

1. In my opinion, the most important criteria for the selection of projects in science and engineering for support are the quality of the investigator and the proposed area of investigation. Relevance to mission needs of an agency are of lesser importance in the support of basic science and engineering. I will illustrate this from my experience over a ten year period, the decade of the '60s, as a member of the scientific advisory panel of the Chemistry Branch of the Air Force Office of Scientific Research (AFOSR). (This agency supports an extramural research program). The program manager was Dr. Amos Horney. At each meeting of the panel (biannual) Dr. Horney opened with his directive to the panel. "Our task is to get the best and the most research with a minimum expenditure of funds for the Air Force. What I want from the panel is to tell me what is the best research; within each subdiscipline give me a rank ordered list; is the proposed budget adequate to do the job or conversely can the proposed budget be cut without reducing the scientific component of the program; what is the size and quality of the total personnel involved in the program; are the facilities adequate. I want to support an investigator's total program; I do not want to support a small component. After that, my responsibility is to see how the program is relevant to the mission of the Air Force".

Let me illustrate how the system performed by citing three programs in chemistry that AFOSR supported during that period and during the '50s.

1. A program by W.F. Libby to study the possibility of archaelogical dating using ^{14}C produced by cosmic radiation. Carbon-14 dating is now a standard tool for dating, including materials of interest to archaeologists and anthropologists. What has been its relevance to the Air Force. The spin off of the program was significant improvements in low level detection of radioactivity. The latter is important in monitoring and detecting nuclear explosions.

2. A program by George C. Pimentel to study high energy chemical systems. Pimentel's work, combined with the pioneering studies of J.C. Polanyi in Toronto on infra-red emission in the reaction of hydrogen and fluorine, led Pimentel to conceive and construct the first chemical lase- a laser whose energy is derived from a chemical reaction carried out in the laser cavity. Chemical lasers are very powerful tools for the study of the rates of chemiical reactions from individual quantum states. But what is its relevance to the Air Force. One of the principal directed energy beams selected for intense development under the Strategic Defense Initiative is the chemical laser, yes the reaction between the rareisotope of hydroge, deuterium, with fluorine.

3. A program by Paul J. Flory on the fundamentals of polymer science. There is hardly an area of polymer science that has not been touched by Flory's work. The Air Force does not manufacture polymers. But it is a big user of polymeric materials in new light weight structures, materials capable of withstanding extremes of temperature and with other important physical and chemical properties, such as electrical insulators, polymeric propellants, just to cite a few examples. Significant improvements have been made in all of these materials, in part, as a result of the understanding Flory has brought to polymer science.

The examples I have cited show the operation of scientific peer review in the selection of quality science for support. I do not suggest that peer review is perfect nor without problems; but it is the best criterion for selection of projects for funding.

The Congress has learned that the Mansfield amendment to the DOD authorization act in the '60s did not serve DOD nor science well. Fortunately, there have been program officers with the vision of Dr. Amos Horney who have worked to minimize the criterion of relevance, even for a mission oriented agency, in the selection of projects for support.

At the individual profram level funding should be at a level consistent with the size of the program and the requirements to carry out the research. How many programs should we fund? This should be determined by the number of quality programs available, by consideration of program areas or sections of the country where stimulation would be desirable, and by our resources. In any case, it is improper for the executive branch, OMB, to counter the decisions of the Congress, as it did recently in attempting to reduce the number of NIH grants to be funded in FY '85.

High Technology firms spend between 2-10% of annual sales on research and development. Most developed countries now spend in the neighborhood of 2 1/2-3 % of annual GNP in government funds on research and development. Such a figure would be appropriate for the United States to maintain its position as a world power and a leader in science and technology.

2. Some students of scientific progress have argued that a steadily increasing science effort and a steadily growing resource level in support of research is necessary to maintain a relatively stable pace of scientific progress, because the cost of research is increasing. The corollary is that level funding will yield a decrease in the rate of scientific discoveries. Do you agree with that analysis, or are there different ways to view the relationship between resources and discoveries? For example, does priority setting within disciplines and among disciplines allow us to optimize the rate of progress?

2. In science, just as in most affairs, experience shows that if you program to maintain steady state you actually fall behind. To maintain steady state as a minimum you must program for growth and progress. Priority setting between disciplines is a very risky exercise, one with which we have little experience. If we had attempted this 40 years ago we would not have the transistor nor the laser and all the science and technology these inventions have spawned. Certainly we would not have the promise of genetic engineering and biotechnology.

Within a discipline it is easier to rank order quality of science and the potential for creating new centers of excellence. These should be our principal guidelines for choice. Relevance should not be a significant consideration in the choice of basic science/engineering programs for support in the academicsector. (I have already touched on this in response to question 1.) Relevance is of importance in selecting areas for support by the private sector, where research and development are related to the company's business activity. Even here the phenomenal success of Bell Laboratories shows the payoff in supporting basic research which is not closely tied to temporal technical requirements, to production or programmatic needs.

3. There is wide agreement that massive scientific advances have been made in the last 40 years. But this has come as the result of the commitment of vast amounts of manpower, funds and other resources. Some have argued that in terms of significant breakthroughs, the volume of high-quality return per unit of investment is actually declining. If that is so, what are the long-term implications for our science policy?

3. I do not support the statement that the volume of high-quality return per unit of investment is declining. We must be careful to define what we are trying to accomplish and to distinguish between apples and oranges. I participated in the scientific advances made 40 years ago. Scientists at any time are limited in the kinds of questions they can ask by the state of knowledge. They are limited in the kinds of investigations they may undertake by the facilities at their disposal. Advances in knowledge, technology and instrumentation have placed unprecedented tools at the disposal of the scientist of the '80s. These make possible studies in areas and at levels undreamed of 40 years ago. Forty years ago the Hale telescope was placed into operation. It has been a major success and continues to be a forefront facility in optical astronomy. But today astronomers are using infra-red telescopes, x-ray telescopes, radio telescopes, and will soon place into operation the Keck telescope and the large space telescope. These instruments will allow us to see farther out into the universe than has been possible heretofore. It will allow us to use information from the entire electromagnetic spectrum, not just visible light, that small region of the spectrum that can penetrate the earth's atmosphere.

Biology has changed from the science which used test tubes, petri dishes, plants and animals; modern biology, the science which has developed from the molecular biology revolution, requires electron microscopes, protein sequencers, high resolution NMR, x-ray crystallography, P-4 installations, etc. All of this instrumentation requires highly trained personnel to operate the equipment.

Space launchers and satellites have brought a wealth of new knowledge about the planets. It has had a major impact on sciences as diverse as astrophysics, atmospheric sciences, geochemistry and geophysics, communications, oceanography, agricultural sciences, weather observation and many more.

Almost every experiment done today is interfaced, one way or another, with a computer. This may be a small controllers, a minicomputer or a supercomputer. The kinds of experiment which can be undertaken today as a result of computer interface and computer analysis were unthinkable 40 years ago.

I could go on and cite examples from almost every field of science and engineering. If we were to decide on what research to undertake and support by the criterion of anticipated return per unit of investment, there would be many research areas which we would either not have undertaken or shut down. One such area currently under discussion from this viewpoint is high energy physics. No serious science policy maker suggests that we should abandon high energy physics because the cost of doing research in this area has grown as we probe ever deeper into the fundamental nature of matter.

4. One observer of science has noted that for individual specialities within science, making discoveries is like picking apples from a tree: Eventually, all the apples are picked, and all the secrets within that specialty are fully discovered. Is that a valid analogy, and if so how can those areas nearing depletion be identified?

4. It is very easy to know when all the apples are picked. The active, vigorous scientist will not wait until 90% of the crop is harvested. He/she will move to another area before that time arrives. There is no need to set out blueprints for programming this movement. The peer review system will clearly detect the time when too many scientists may be looking for the last apple. At the same time, if a scientist with a proven tract record feels there is more fruit to be harvested when fashions suggest that 90% of the crop is in, we should him/her to risk his/her future on his/her judgment.

From time to time a new discovery in another discipline, a new research tool, or the fresh insight brought by a new young investigator has the capacity of rejuvenating a "picked over" discipline and cause it to flourish. I cited one such example, atomic physics, in my prepared testimony to this Committee.

5. Dr. Bentley Glass, in his Presidential address to the A.A.A.S. in 1971, said "What remains to be learned may indeed dwarf the imagination . . . We are like the explorers of a great continent who have penetrated to its margins in most points of the compass and have mapped the major mountain chains and rivers. There are still innumerable details to fill in, but the endless horizon no longer exists." Since Vannevar Bush spoke of the "endless frontier" in 1945, has science, in fact, experienced, like the American West in the 19th century, the end of the frontier?

5. Unlike the opening of a continent, science does not have a frontier boundary. Each new major discovery opens up new areas for investigation. In fact, one of Alfred Nobel's objectives in establishing the Nobel prizes was to recognize the most outstanding discovery each year in chemistry, physics, and the life sciences which also had the potential for opening up entire new areas of investigation. While the Nobel prizes tend to lag between the time of discovery and recognition, they have consistently lived up to Nobel's expectation. There is no sign that the process is at an end.

I have reviewed my response with my colleague, Dr. Bentley Glass, who finds himself in complete agreement with my response. His statement in his 1971 Presidential address to the AAAS should be understood within the context of a number of issues raied in these hearings. Progress within a scientific discipline follows a sigmoid type curve. There is an induction period, a period of slow progress; some breakthrough leads to a period of exponential growth, which is then followed by a slow down in the rate of growth. During this last phase, new results, which may be very important, come at a slower rate. This phase may actually serve as the induction period for the next phase of rapid growth within the discipline.

6. In your view, would one of the goals of government science policy be to achieve and maintain, as a matter of national prestige, U.S. leadership across the spectrum of science, or should we share or yield leadership in some areas of science to other countries?

6. Our goal is to be preeminent in all areas where we have scientific quality and an adequate supply of manpower. We should work toward international cooperation since this benefits all of science. We should stimulate interest in those areas of good science where we may lag behind other major countries. Our strategy should not be "have we made a clean sweep of the gold medals?" Nor can we entertain an attitude that we cannot afford to be in the forefront in all areas of science. If we are not in the forefront, it should be because of the intersts of our scientists, the number of our scientists working in the area, or the quality of our scientists working in the area.Our national policies and priorities should ensure that we have the educational system and a set of values that will continue to producegenerations of scientists comparable to and better than those that we produced, and attracted from abroad, over the past 40 years.

7. In your view, have the goals and objectives of federal support for science changed over the years since they were outlined in the Bush Report 40 years ago? As you look beyond the current studies and science budgets for the next few years, what changes or adjustments in our goals and objectives do you foresee for the decades after year 2000?

7. The goals and objectives outlined in the Bush report were far sighted; they have served us well. I see no reason to change those goals and objectives.

8. The Chairman of the National Academy's Space Science Board recently noted that there are no scientific criteria we can develop for science as a whole. He said that "[w]e are experts at setting priorities within any one field of science. The astronomer, for example, finds it difficult to judge impartially the value of research in the life sciences. The ultimate judgement about priorities are made adequately by the present method of relying on a complex democratic process to make essential political decisions." What is your view of the roles of scientists and politicians in making decisions about scientific priorities?

8. I fully support the statement of Dr. Thomas Donahue, Chairman of the NRC Space Science Board, with regard to the limits of our abilities as scientists to set priorities. The scientist can tell thosed charged with setting national policies and with the expenditure of public funds what is good science. Our representatives n the Congress and the members of the Executive branch have the responsibilities of establishing our national objectives and priorities. They should use information and guidance from the scientific community in arriving at their decisions. They must consider questions of national budgets, the needs of the government itself for scientific knowledge, the need to develop excellence in scientific research across the nation, the impact of government funding on local and regional economies, etc. A deliberative process, which provides for input from the scientific community a appropriate stages is of national interest.

9. Most studies of science and most agency budgets for science are future oriented. They speak of future opportunities, future projects, and future results. Retrospective discussions are limited to anecdotal cases of successes, while little has been done to look carefully at entire programs and the ratio of those which lead to technological successes and those which do not, however measured. Why should not more such comprehensive evaluations of past programs be done?

9. Retrospective comprehensive evaluations have the potential of providing guidance in planning future programs. For instance, such studies might show the relative effectiveness of programs like RANN compared with unstructured grants to individual investigators. In our pluralistic mode of support, which is a great strength in the United States system for the federal support of science, it could highlight those agencies which have particularly outstanding and effective programs. This might lead to changes in the policies under which some agencies operate; some program officers may find new methods to improve the operation of their programs. Had such a process been in place 35 years ago we wouldnot have gone for 35 years before AEC, ERDA and DOE changed its mandated policy for funding extramural research from annual contracts to multiyear grants. Annual grants have been extremely wasteful of time and resources of both the principal investigator and the program managers.

10. Some 20 years ago, Dr. Harvey Brooks, who then chaired a comparable committee, noted in a report to this Committee and quoting an OECD report, that "most countries have better statistics on poultry production than on the activities of their scientists and engineers." Your present report does not discuss that issue, but neither does it employ, to any extent, quantitative analysis for policy formulation, by more expensive use of such data. What, in your view, are the prospects for such expanded use of statistical analysis in the formulation of U.S. science policy?

10. "The Outlook for Science and Technology 1985" is not encyclopaedic. Nor is it meant to be the only source of information available to the House S&T Committee. NSF has an entire division devoted to program analysis, statistics on numbers of scientists, expenditures by all agencies for each field of science. Our report has utilized those data to highlight for the Congress issues related to scientific and engineering personnel. We have done this on a timely basis; subsequent analyses have indeed reached conclusions identical with those in the "Outlook" report.

The "Outlook" also discusses issues related to facilities and instrumentation. These are based in part on statistical studies carried out by the National Research Council, the Commission on Research of AAU and other sources. In the interest of brevity we have not expanded the "Outlook" to include extensive tables of numbers, ages of scientists, costs of all types of instruments by discipline, etc. Such detailed tables are available from the National Research Council, the National Science Foundation and others.

I would anticipate that future issues of "The Outlook for Science and Technology" will highlight for the Congress other issues based on statistical analyses. Certainly the maintenance of an adequate data base is essential for the efficient operation of our national science program.

QUESTIONS FOR THE RECORD

Dr. Floyd E. Bloom

1. In your prepared testimony, you noted that if a "recess" in support for science should occur, there is "the likely possibility that the important discoveries and their applications and economic gains will be won elsewhere." Would you define a recess in support as slower growth, level funding, or a funding reduction? and, since publication of research results disregards national boundaries, can U.S. technology and medicine not benefit fully from discoveries made in other countries?

1. In the context of the response I made, a "recess" connotes reduced numbers of dollars available per approved project. A stable level of funding, buffered from inflation, would permit scientists and science administrators a more accurate estimate of opportunities on which to base proposals and commitments for future activities, and could eliminate the hot-cold alterations that have been in place over the past few years.

2. Your report begins with two highly important and crucial questions: "Of current and potential endeavors in science and engineering, what criteria should be used in selecting projects to be funded? What determines appropriate funding levels?" Can you offer any suggestions about what the answers to those questions are, or how we can go about developing the answers?

2. This complex matter is one that our COSEPUP is seriously examining, and which our Chairman, Dr. Omenn, will I'm certain be willing to share with your staff, when a few viable alternative ideas have been collected.

3. Some students of scientific progress have argued that a steadily increasing science effort and a steadily growing resource level in support of research is necessary to maintain a relatively stable pace of scientific progress, because the cost of research is increasing. The corollary is that level funding will yield a decrease in the rate of scientific discoveries. Do you agree with that analysis, or are there different ways to view the relationship between resources and discoveries? For example, does priority setting within disciplines and among disciplines allow us to optimize the rate of progress?

3. Within the viewpoint expressed in my response to question 1, level funding would not necessarily preclude continued progress or discovery, but would (in my view again) force scientists and science administrators to make decisions that they now prefer to leave to political demands and to make priority decisions within and across fields. Since there are no known equitable ways to do that, mistakes will be made, and some discoveries will undoubtedly be delayed or not made at all. This is a part of the equation that leads to the concept that within available national funds, science research funding should be level, rather than being made available as needed to reach the outcomes that are desired (better health, better technologies, for better products for a growing economy, etc).

4. There is wide agreement that massive scientific advances have been made in the last 40 years. But this has come as the result of the commitment of vast amounts of manpower, funds and other resources. Some have argued that in terms of significant breakthroughs, the volume of high-quality return per unit of investment is actually declining. If that is so, what are the long-term implications for our science policy?

4. I cannot accept the premise on which this question is based. Important, significant breakthroughs are occurring within my fields of science (medical biology generally, and brain sciences, specifically) unprecedented rates, with enormous benefits. This tremendous harvesting of important practical knowledge does indeed represent decades of enormous committment. However, nature does not yield her secrets easily. New tools and discoveries in related fields have all converged to make the present decade the richest in discoveries ever known within biology, and the yields are likely to pay dividends for broader discoveries and control of serious disorders for decades to come.

5. One observer of science has noted that for individual specialities within science, making discoveries is like picking apples from a tree: Eventually, all the apples are picked, and all the secrets within that specialty are fully discovered. Is that a valid analogy, and if so how can those areas nearing depletion be identified?

5. Again, I respectfully disagree with this "observer". Science is a means to ask questions about man's presence and future in a world of unpredictable physical and social forces, limited life spans, and limited, and unequally distributed resources. When scientific questions are phrased properly and viewed effectively, good answers lead to better questions, not to extinction of curiosity and a final version of "truth". The most effective existing means of peer priority setting are defining those parts of the knowledge orchard where future progress soon are unlikely; these peers vote with their feet in moving to new places where effective participation is more feasible. In my view, the issue underlying this question is not a major problem.

6. Dr. Bentley Glass, in his Presidential address to the A.A.A.S. in 1971, said "What remains to be learned may indeed dwarf the imagination . . . We are like the explorers of a great continent who have penetrated to its margins in most points of the compass and have mapped the major mountain chains and rivers. There are still innumerable details to fill in, but the endless horizon no longer exists." Since Vannevar Bush spoke of the "endless frontier" in 1945, has science, in fact, experienced, like the American West in the 19th century, the end of the frontier?

6. I believe my answer to question 5 subsumes the response here, which in brief is NO.

7. In your view, have the goals and objectives of federal support for science changed over the years since they were outlined in the Bush report 40 years ago? As you look beyond the current studies and science budgets for the next few years, what changes or adjustments in our goals and objectives do you foresee for the decades after year 2000?

7. I regret that I do not have the experience to provide a solid answer to this concern with regard to the Bush report's original directions. With regard to future directions after the year 2000, I see our main objectives and goals to require continued peaceful co-existence with the capacity to devote greater proportions of available national resources to improvements of disease diagnosis, and prevention, and to a general concern for the maintenance of the existing resources of our planet.

8. In your view, would one of the goals of government science policy be to achieve and maintain, as a matter of national prestige, U.S. leadership across the spectrum of science, or should we share or yield leadership in some areas of science to other countries?

8. In my view, US leadership should be sought across the board of science and technology. This point, which I and the COSEPUP spoke to specifically in the Fourth Five Year Outlook Report discussions, is based on the view that one cannot choose to be number 2 or less in a field that may be important tomorrow without relinquishing the capacity for shared and spreading discovery that emerges from a broadly based pursuit of excellence in discovery and application.

9. The Chairman of the National Academy's Space Science Board recently noted that there are no scientific criteria we can develop for science as a whole. He said that "[w]e are experts at setting priorities within any one field of science. The astronomer, for example, finds it difficult to judge impartially the value of research in the life sciences. The ultimate judgement about priorities are made adequately by the present method of relying on a complex democratic process to make essential political decisions." What is your view of the roles of scientists and politicians in making decisions about scientific priorities?

9. While I agree in general with the view stated, I do feel that scientists may not have adequately provided in the past the kinds of information and perspective that could have been helpful at times when elected officials will need to make critical political judgements. For that reason, I am always pleased to have had the opportunity to work in presenting those items of information to contribute to the continuing dialog necessary to proceed as wisely as is possible in a democracy.

10. Some 20 years ago, Dr. Harvey Brooks, who then chaired a comparable committee, noted in a report to this Committee, and quoting an OECD report, that "most countries have better statistics on poultry production than on the activities of their scientists and engineers." Your present report does not discuss that issue, but neither does it employ, to any exent, quantitative analysis for policy formulation, by more expensive use of such data. What, in your view, are the prospects for such expanded use of statistical analysis in the formulation of U.S. science policy?

10. This question requires more information than I have, but to the degree that better data on the numbers of scientists in research, and in the pipeline of training to enter research or education or industry, can provide useful details, such data bases are being developed. While these numbers may be helpful in completing a general picture of that state of science and technology at any one time, the essence of discovery and progress is unlikely to be reflected in these numbers. Ideas emerge from critical minds in critical masses that vary from time to time. A poultry farm is a biological factory for the production of eggs and chicken products; the right ratios of space, temperatures and lighting can indeed enhance chicken productivity. The rules for fostering creativity and industrial application are not now knowable, and I do not feel that statistics on the numbers of thinkers and thinking sites will help provide them.

QUESTIONS FOR THE RECORD

Dr. Herbert A. Simon

1. Your report begins with two highly important and crucial questions: "Of current and potential endeavors in science and engineering, what criteria should be used in selecting projects to be funded? What determines appropriate funding levels?" Can you offer any suggestions about what the answers to those questions are, or how we can go about developing the answers?

1. Criteria for Funding. It is hard to throw any new light on this basic and unavoidable question. Perhaps the only answer is "Listen and weigh" -- that the Congress will have to continue its policy of letting scientists in the various disciplines make the cases for research in their fields, and then weigh their claims for the contributions of their fields to basic knowledge and thence to the solution of national problems. As we work down the hierarchy to narrower and narrower domains of science, of course the scientists themselves can be of greater help in advising on priorities, as the astronomers have done, and as we regularly do at the grant level through devices like peer review. This present system, while logically unsatisfactory, is perhaps not all bad -- it has given us lots of good science and valuable new ideas and technology relevant to social problems.

In weighing the social value of supporting science, of course it is necessary not only to think in terms of balance between sciences (a very difficult concept to implement), but also of the relation between research and the social needs it is serving. In a society that spends a tenth of its income for medical services, it is easy to see how a very large budget for biological and medical research can be justified. In a society that spends a comparable amount on education, one could think that a roughly similar research effort in that domain would pay off also. I am not suggesting a hard-and-fast rule of "taxing" each segment of our society at some fixed rate to support research, but that this is at least one way of thinking about the range and magnitude of social needs being served by research in a particular domain.

2. Some students of scientific progress have argued that a steadily increasing science effort and a steadily growing resource level in support of research is necessary to maintain a relatively stable pace of scientific progress, because the cost of research is increasing. The corollary is that level funding will yield a decrease in the rate of scientific discoveries. Do you agree with that analysis, or are there different ways to view the relationship between resources and discoveries? For example, does priority setting within disciplines and among disciplines allow us to optimize the rate of progress?

2. Rate of progress. There is no persuasive evidence that we have reached a point of diminishing returns in science as a whole; but there are continual changes in priorities -- in the domains of science in which the "hot" areas lie. The attractiveness of areas to young scientists is not a bad (if very rough) measure of where the action is. In the long run, aggregate expenditures for science can only grow as national income grows, and the criterion suggested in the answer to Question 1 -- relating research expenditures to the magnitude of the sector of society being served -- gives at least a rough and ready, if not deeply defensible, measure of justifiable research levels.

3. There is wide agreement that massive scientific advances have been made in the last 40 years. But this has come as the result of the commitment of vast amounts of manpower, funds and other resources. Some have argued that in terms of significant breakthroughs, the volume of high-quality return per unit of investment is actually declining. If that is so, what are the long-term implications for our science policy?

3. Diminishing returns. I cannot agree with the premise of the question: that scientific research has passed the point of diminishing returns. On the contrary, very recent progress in such domains as solid state physics, biotechnology, and cognitive science is opening up to us new approaches to a whole range of problems of first magnitude. We have the human resources in our country to maintain first-rate scientific activity, at current levels in relation to income and population, into the foreseeable future.

4. One observer of science has noted that for individual specialities within science, making discoveries is like picking apples from a tree: Eventually, all the apples are picked, and all the secrets within that specialty are fully discovered. Is that a valid analogy, and if so how can those areas nearing depletion be identified?

4. Areas nearing depletion. While there certainly are lags in identifying areas nearing depletion, scientists are continually redeploying their resources to tackle the new promising areas and to abandon those that have been worked out. The rapid, nearly spontaneous, growth of areas like molecular biology, computer science, and cognitive science are examples of this. Scientists are highly motivated to search out the directions where there is pay dirt because their own careers are based on the success of their research. Hence, the scientists within a discipline can usually be relied on for fairly accurate judgments of the directions in which progress will be made, although there is necessarily some time lag in these judgments.

5. Dr. Bentley Glass, in his Presidential address to the A.A.A.S. in 1971, said "What remains to be learned may indeed dwarf the imagination . . . We are like the explorers of a great continent who have penetrated to its margins in most points of the compass and have mapped the major mountain chains and rivers. There are still innumerable details to fill in, but the endless horizon no longer exists." Since Vannevar Bush spoke of the "endless frontier" in 1945, has science, in fact, experienced, like the American West in the 19th century, the end of the frontier?

5. Filling in details. It would be interesting to make a collection of the statements that scientists have published in the past announcing that the last frontier had been reached. Such statements were especially common around the end of the last century, when Newton and Maxwell had defined the boundaries. But with 1900 came Planck with the quantum, and in 1905 Einstein with relativity, and we realized once again that we hadn't been near the frontiers.

Do we think today that quarks will still be the "fundamental particles" fifty years from now, or that we will then still be in ignorance of the physico-chemical basis of human memory? Perhaps, but don't bet any money on it. Of course one person's revolutionary advance may be another person's "detail." As long as human ignorance bars us from dealing successfully with human problems -- physical, biological, or behavioral -- of first importance, there remain plenty of frontiers for science.

6. **In your view, would one of the goals of government science policy be to achieve and maintain, as a matter of national prestige, U.S. leadership across the spectrum of science, or should we share or yield leadership in some areas of science to other countries?**

6. U. S. leadership. National prestige can only be a secondary objective. The primary objective is to create new knowledge that meets human needs, and to do this in cooperation with scientists around the world. We are in the fortunate position, in the U.S., of being able to devote vast resources, human and financial, to science, hence of often being in the vanguard. We should seek to maintain a position in science that is commensurate with our resources. That does not mean that we will always be first in every area. We should compete with ignorance, not with nations.

7. **In your view, have the goals and objectives of federal support for science changed over the years since they were outlined in the Bush Report 40 years ago? As you look beyond the current studies and science budgets for the next few years, what changes or adjustments in our goals and objectives do you foresee for the decades after year 2000?**

7. Changes in Goals and Objectives. I must answer this question from the special perspective of a behavioral scientist. Since the Bush report was published, there has been a growing recognition of the role the social and behavioral sciences can and should have in the total spectrum of science. These sciences have been incorporated in the National Academy of Science, and have gradually secured broadened support from the National Science Foundation and other funding agencies; although, as we are well aware, there are still ups and downs in the solidity of the support for these sciences. Given their importance to our society -- in deepening our understanding of ourselves, in helping us to improve the processes of management and of education, of designing better institutions for criminal justice and for meeting the needs of all sectors of our society -- I expect that the present long-term trend, toward expanded research effort in the social and behavioral sciences, will continue and even accelerate.

8. The Chairman of the National Academy's Space Science Board recently noted that there are no scientific criteria we can develop for science as a whole. He said that "[w]e are experts at setting priorities within any one field of science. The astronomer, for example, finds it difficult to judge impartially the value of research in the life sciences. The ultimate judgement about priorities are made adequately by the present method of relying on a complex democratic process to make essential political decisions." What is your view of the roles of scientists and politicians in making decisions about scientific priorities?

8. Scientific priorities. The answer to this question is related to the answer to Question 1. Decisions about priorities involve both technical issues, on which scientists are perhaps best informed, and issues of social value, on which politicians should decide. This is, of course, oversimplistic; but the line between fact and value still gives us a rough guideline of where it is useful to bring scientists into the decision process: it is most obviously helpful in judging priorities within a single field; much less helpful in advising on allocations between broad areas that are quite distinct, and which also commonly serve quite different social needs.

But what is, of course, needed most of all is communication and cooperation between scientists and politicians in trying to reach commonsensical answers to these difficult questions of resource allocation. We can be skeptical that there are unique correct answers, yet continue to strive towards answers that are not unreasonable -- answers that make heavy, but not unconscionable demands on social resources, and that point efforts in directions that will relate them clearly to important social needs.

9. **Most studies of science and most agency budgets for science are future oriented. They speak of future opportunities, future projects, and future results. Retrospective discussions are limited to anecdotal cases of successes, while little has been done to look carefully at entire programs and the ratio of those which lead to technological successes and those which do not, however measured. Why should not more such comprehensive evaluations of past programs be done?**

9. Evaluation. Retrospective evaluation can give us useful insights about whether we have been behaving sensibly in the ways in which we go about funding and managing science. Several informative evaluation studies have been made in the past, although they can, perhaps, be accused of being anecdotal. What we should not expect is that retrospective studies can give quantitative answers, or that we can always trace out the cause-and-effect of whether particular areas of basic research have produced "technological successes" in subsequent application. Many factors besides the quality of basic research determine whether it is successfully converted into technology -- including the willingness and ability of the institutions that manage technology to accept it. (The fact that the Japanese have made better use of our basic research on methods of industrial quality control than we have is an interesting example to ponder.)

Yes, we should have more retrospective evaluation, but comprehensive, aggregative statistical studies, will be far less revealing than careful case studies of specific areas of science and specific developments. And we should study both the successes and the failures.

10. Some 20 years ago, Dr. Harvey Brooks, who then chaired a comparable committee, noted in a report to this Committee, and quoting an OECD report, that "most countries have better statistics on poultry production than on the activities of their scientists and engineers." Your present report does not discuss that issue, but neither does it employ, to any exent, quantitative analysis for policy formulation, by more expensive use of such data. What, in your view, are the prospects for such expanded use of statistical analysis in the formulation of U.S. science policy?

10. Statistical data. The answer to Question 9 reveals my skepticism that an expanded use of statistical analysis is the best way to improve our understanding of the scientific enterprise and to guide it better. Careful research on the processes of science and its institutions, and periodic reviews (of the sort that have been made on a number of occasions under the auspices of the Academy's Committee on Science, Engineering, and Public Policy) of the progress and problems of particular disciplines will be more productive than statistical studies, however sophisticated.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

APPENDIX 2

Additional Materials for the Record

The Outlook for Science and Technology 1985

Committee on Science,
Engineering, and Public Policy

National Academy of Sciences
National Academy of Engineering
Institute of Medicine

NOTICE: The National Academy of Sciences was established in 1863 by Act of Congress as a private, nonprofit, self-governing membership corporation for the furtherance of science and technology for the general welfare. The terms of its charter require the National Academy of Sciences to advise the federal government upon request within its fields of competence. Under this corporate charter, the National Academy of Engineering and the Institute of Medicine were established in 1964 and 1970, respectively.

The Committee on Science, Engineering, and Public Policy is a joint committee of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. It includes members of the councils of all three bodies.

This work was supported by the National Science Foundation under Grant PSP-7913104.

Library of Congress Catalog Card Number 85-60421

Available for $3.00 per copy (prepaid only) from

National Academy Press
2101 Constitution Ave., NW
Washington, DC 20418

Printed in the United States of America

Preface

The 1985 Outlook for Science and Technology is the fourth report in a series originally mandated by the National Science and Technology Policy, Organization, and Priorities Act of 1976. Subsequent to the passage of that Act, a reorganization of the Executive Office of the President gave the National Science Foundation responsibility for preparing the Outlooks. In turn, the Foundation asked the National Academy of Sciences to assist. Since 1982, the Outlooks have been prepared by the Committee on Science, Engineering, and Public Policy (COSEPUP), a joint unit of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The purposes of these reports are to describe and discuss:

- current and emerging problems of national significance that are identified through scientific research or in which scientific or technical considerations are of major importance; and
- opportunities for the use of new and existing scientific and

technological capabilities that can help to resolve these problems and impediments to the effective use of these capabilities.

Issues presented in this Outlook were obtained from several sources, including discussion within COSEPUP and from research briefings prepared for the President's Office of Science and Technology Policy (OSTP), the National Science Foundation, and other federal agencies. These briefings summarized research opportunities with the potential for high scientific returns in a number of fields, identified by OSTP after consultation with COSEPUP. Twenty-one briefings have been presented in the last three years on subjects as diverse as agriculture, astronomy, atherosclerosis, catalysis and other topics in chemistry, computer-aided manufacturing, information technology in precollege education, neuroscience, immunology, and cognitive science and artificial intelligence.

In this report, we summarize the nine subjects that formed the bases of the 1984 research briefings. Subjects of comparable scientific merit could have been selected. However, as with past briefings, the topics selected were those for which near-term federal actions were considered likely.

Other suggestions for this Outlook came from the major units of the National Research Council and from Councilors of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The committee is grateful to all of these contributors and especially to the members of its subcommittee responsible for preparation of the 1985 Outlook: Jacob Bigeleisen, of the State University of New York at Stony Brook (chairman); Floyd E. Bloom, of the Scripps Clinic and Research Foundation; Emilio Q. Daddario, of Wilkes, Artis, Hedrick, and Lane; and Edward A. Mason, of the Standard Oil Company (Indiana). The substantial contributions by Norman Metzger in preparing this report are acknowledged.

LEON T. SILVER
Chairman
Committee on Science, Engineering,
and Public Policy

Committee on Science, Engineering, and Public Policy

Subcommittee on the Outlook for Science and Technology

JACOB BIGELEISEN, Leading Professor, Department of Chemistry, State University of New York at Stony Brook (*Chairman*)
FLOYD E. BLOOM, Director and Member, Division of Preclinical Neuroscience and Endocrinology, Scripps Clinic and Research Foundation, La Jolla, California
EMILIO Q. DADDARIO, Wilkes, Artis, Hedrick, and Lane, Attorneys at Law, Washington, D.C.
EDWARD A. MASON, Vice President, Research, Standard Oil Company (Indiana), Naperville, Illinois

Staff

NORMAN METZGER, Staff Officer
AUDREY PENDERGAST, Associate Editor
GERRY KASARDA, Administrative Assistant

Committee on Science, Engineering, and Public Policy

Leon T. Silver, William M. Keck Foundation Professor of Geology, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California (*Chairman*)

Linda H. Aiken,* Vice President for Research, The Robert Wood Johnson Foundation, Princeton, New Jersey

Richard C. Atkinson, Chancellor, University of California, San Diego

Jacob Bigeleisen, Leading Professor, Department of Chemistry, State University of New York at Stony Brook

Floyd E. Bloom, Director and Member, Division of Preclinical Neuroscience and Endocrinology, Scripps Clinic and Research Foundation, La Jolla, California

W. Dale Compton, Vice President, Research, Ford Motor Company, Dearborn, Michigan

Emilio Q. Daddario, Wilkes, Artis, Hedrick, and Lane, Attorneys at Law, Washington, D.C.

Gerald P. Dinneen, Vice President, Science and Technology, Honeywell, Inc., Minneapolis, Minnesota

Alexander Flax, President Emeritus, Institute for Defense Analyses, Alexandria, Virginia

Zvi Griliches, Professor, Department of Economics, Harvard University, Cambridge, Massachusetts

Philip Leder, John Emory Andrus Professor and Chairman, Department of Genetics, Harvard Medical School, Cambridge, Massachusetts

Gardner Lindzey, President and Director, Center for Advanced Study in the Behavioral Sciences, Palo Alto, California

---

*Term expired December 31, 1984.

EDWARD A. MASON, Vice President, Research, Standard Oil Company (Indiana), Naperville, Illinois

DANIEL NATHANS, Professor, Department of Molecular Biology and Genetics, The Johns Hopkins University School of Medicine, Baltimore, Maryland

GILBERT S. OMENN, Dean, School of Public Health and Community Medicine, University of Washington, Seattle, Washington

HERBERT A. SIMON, Richard King Mellon University Professor of Computer Science and Psychology, Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania

F. KARL WILLENBROCK, Cecil H. Green Professor of Engineering, School of Engineering and Applied Science, Southern Methodist University, Dallas, Texas

Ex Officio Members

FRANK PRESS, President, National Academy of Sciences
ROBERT M. WHITE, President, National Academy of Engineering
FREDERICK C. ROBBINS, President, Institute of Medicine

COSEPUP Staff

ALLAN R. HOFFMAN, Executive Director
BARBARA A. CANDLAND, Administrative Assistant
JOANNA M. MASTANTUONO, Senior Secretary

Contents

OVERVIEW 1

I. RECENT PROGRESS IN SCIENCE AND TECHNOLOGY 5
  Oncogenes 5
  Atherosclerosis 6
  Parasitism 8
  Chemical and Process Engineering for Biotechnology 9
  Advanced Polymeric Composites 10
  Supercomputer Architectures 11
  Information Technology in Precollege Education 12
  Opportunities in Physics 13
  Solar-Terrestrial Plasma Physics 14

II. AN OUTLINE OF SELECTED ISSUES 17
International Competition in Science and Technology 18
Supercomputers, 18; Biochemical Engineering, 20; Advanced Polymeric Composites, 22
Issues for the Congress, 22
Scientific and Engineering Personnel 23
Starting a Research Career, 23; Clinicians in Research, 24; Possible Shortages, 24; Doctorates for Non-U.S. Citizens, 25
Issues for the Congress, 27
Cooperative Work Across Disciplines 28
The Special Case of Agriculture, 29
Issues for the Congress, 30
Research and Transportation 30
Highways, 31
Issues for the Congress, 32
Facilities and Instrumentation 33
Large-Scale Facilities, 33; Medium-Scale Instrumentation, 34; Planning New Facilities, 37
Issues for the Congress, 38
Issues in Genetic Engineering 40
Issues for the Congress, 41
Issues in Human Biology 41
In Vitro Fertilization and Implantation of Human Eggs, 42; Genetic Therapy of Human Cells, 42
Issues for the Congress, 44
Scientific Communication, Technology Transfer, and National Security 44
The Department of Defense, 45; The Department of Commerce, 45
Issues for the Congress, 47
Global Atmospheric Effects of Nuclear Explosions 48
Issues for the Congress, 49
Final Comment 49

The Outlook for Science and Technology 1985

Overview

That basic research is vital to the nation's future is now an axiom of national policy. In their statements and in their budget proposals, both the Administration and the Congress agree that progress in fundamental science and engineering is essential to the economy of a modern technological society. This position is taken by other industrialized nations and is reflected in their rising investments in research.

Granting this acceptance of the value of an aggressive national program in basic research, there are limits to the resources available and hard choices to be made in budget priorities. As in other federal endeavors, those concerned with the levels of support for science and technology face difficult decisions:

- Of current and potential endeavors in science and engineering, what criteria should be used in selecting projects to be funded? What determines appropriate funding levels?
- How can the needs for large facilities—whether refurbishing those that have become outdated or constructing new ones—long evident in many fields of science and engineering be satisfied

without sharp perturbations in the federal budget for research and development?

- Are mechanisms and support available to fuel shifting research patterns, such as the heightened interactions among disciplines and between academic and industrial researchers?

To state these questions is not to assert that they are readily answerable. Progress toward answering them must come by exposing the problems and the opportunities.

The Outlook begins with brief glimpses of nine vibrant fields of science and technology. The status of each field was analyzed intensively in a series of research briefings prepared in 1984 by the Committee on Science, Engineering, and Public Policy and was presented to the President's Office of Science and Technology Policy, the National Science Foundation, and other federal agencies. In actuality, the menu of the particular fields chosen is not crucial; any other roster of significant topics would yield the same generalizations:

- *American science and technology continue to be amazingly fertile.* Whatever the particular field, the outcome is the same: new, sometimes startling, discoveries and unexpected progress in dealing with wider problems emerge, from the understanding of cancer and atherosclerosis to creating a new generation of electronic and optical devices. The same story applies to fields at the interface of science and technology—computer architecture, advanced composites, and engineering involving biotechnology; in these latter fields, the enormous interdependence of basic science and engineering knowledge is even more evident.
- *Research advances depend increasingly on the combined efforts of many disciplines.* Progress on oncogenes, for example, came not only from work on the cancer problem itself, but also from other fields: virology, molecular and cellular biology, pharmacology, and biochemistry.
- *Progress in basic science and technology is happening at an incredibly fast pace.* Ten years ago, Part I of this Outlook could not have been written: oncogenes were speculations at best; atherosclerosis remained a deeply puzzling affliction. The laser and integrated circuit chips were recent arrivals, while such contemporary techniques as molecular beam epitaxy and a host of

new spectroscopies such as femtosecond spectroscopy did not seem realizable. And, new technologies have arrived, from biotechnology to the imminent debut of computer architectures embodying the synchronous operation of several hundred processors.

- *The gap between new knowledge and its application has narrowed.* This report points out two examples of prompt application: the enormous industry upwelling from fundamental discoveries in gene manipulation and a totally new class of semiconductors arising from basic and quite recent advances in surface science.

The linked themes then are that: science and technology are being deeply refreshed by a torrent of new discoveries; in many fields, these discoveries are coming at an ever faster pace, and they are being translated rapidly into wider use.

In short, the American public is reaping the rewards for its continuing support of fundamental research. And it is reaping the rewards of having made the university the home of basic research, an achievement virtually unique among nations. In doing so, it has created a rich atmosphere of learning and discovery, of teaching and research, of unrelenting questioning as teachers and their students combine to attack the puzzles of nature.

The issues discussed in this Outlook deal with the present and future health of the research system and its principal components—the universities, industry, and government, and their relationships. These issues arise from several directions and include the following:

- *The international competitive strength of the United States.* That issue is a tangled one, although one certain thread is the standing of various fields of science and technology, especially those currently acknowledged as economically vital. This Outlook considers that issue, by examining the competitive status of several fields and suggesting some responses fitted to the U.S. system.
- *Scientific and engineering personnel.* The Outlook addresses several facets, among them counterincentives for promising students to enter research careers, illustrated by the difficulties faced by young investigators, and the role of foreign students in U.S. advanced education.

- *Facilities and instrumentation*. The almost desperate need of research universities to rejuvenate existing technical facilities and to build new ones, as well as the paucity of state-of-the-art instrumentation in many academic laboratories, is recognized increasingly—in the Congress and the Administration. This Outlook attempts to sort out the various needs. Recent episodes in which universities have obtained funding for facilities using the direct appropriation mode are also discussed.
- *Open scientific and technological communication and national security*. While the tensions inherent in this issue are now quite public, they continue to be unresolved. Considerable uncertainty burdens a number of research fields with military import, while restrictions, existing or prospective, on the flow of technical data have hindered or may hinder industrial progress, especially in the application of advanced technologies.

Other matters also are treated in this report, from human biology to research on the uncertainties embedded in the nuclear winter concept. Whether new or continuing, these are some of the issues that come to the fore in considering systematically the status and outlook for American science and technology.

This Outlook has two major parts.

Part I is a precis of recent progress in certain fields of science and technology, taken from nine research briefings prepared in 1984 by the Committee on Science, Engineering, and Public Policy.*

Part II summarizes several national issues that can be inferred from the fields covered in these briefings and from other areas of science and technology.

---

*The full text of the briefings summarized here has been published, as *Research Briefings 1984* and is available from the National Academy Press, 2101 Constitution Avenue, NW, Washington, DC 20418.

PART I

Recent Progress in Science and Technology

Oncogenes

How is a normal cell transformed into a cancer cell? How can diverse agents, from chemicals to radiation to viruses, cause that transformation? The answers to these questions require an understanding of the molecular changes that propel a normal cell into malignancy.

This understanding has been emerging in the last 10 years from intersecting work in several subfields of biology, among them cellular and molecular biology, pharmacology, and biochemistry. As a result, key aspects of cancer can now be described in molecular terms: normal genes that control cell growth become slightly modified. These modified genes then encode proteins capable of changing a normal cell into a cancer cell; hence, those genes are called oncogenes.

Oncogenes were first discovered through studies of animal cells infected by viruses, including the Rous sarcoma virus that causes cancers in chickens. The virus converts certain normal animal genes into potent oncogenes.

Human oncogenes were discovered by inserting DNA segments from human cancer cells into normal cells in culture. The specific DNA sequences responsible for transforming these recipient cells into cancerous cells—human oncogenes—are closely related both to normal human genes and to viral oncogenes. Active oncogenes have been demonstrated in a variety of human cancers.

Knowledge of the structures of oncogenes, their relation to chromosomal abnormalities seen in malignancies, the proteins they encode, and the intriguing relation of some oncogenes to growth factors observed in hormonal tissue repair has expanded enormously in recent years. However, exactly how oncogenes act, the functions of the proteins they encode, and the nature of their activation by chemical carcinogens, viruses, radiation, and other agents are still unclear.

Cancers are diverse; they have neither a single cause nor a single cure. Further, the transformation of normal cells into malignant ones includes many steps. Among them, the activation of oncogenes is an important, perhaps necessary step, but not the only one. While efforts to prevent cancers can be directed against any of these critical stages, the discovery of some 20 human oncogenes has expanded possibilities for the treatment and prevention of cancers. There could be drugs to block the action of oncogene proteins; or immunologic agents, including antibodies, that would recognize and destroy cells carrying oncogene proteins on their surfaces; or agents to block cellular receptors that enhance the growth of malignant cells. The diagnosis of cancers also may be improved by identifying oncogenes activated by an environmental or other agent.

Atherosclerosis

Atherosclerosis causes heart attacks and strokes and accounts for half of all of the deaths in the United States. In this disease, the flow of blood through the arteries is obstructed by plaques that have formed on the arterial walls. Eventually, a blood clot develops and obstructs the artery, blocking the flow of blood to the heart muscle or the brain.

This disease is the product of a complex interplay between components of the blood and the cells that line the interior walls of the blood vessels. Through interdisciplinary efforts drawing upon molecular and cellular biology, physics, chemistry, and genetics, that interplay has become better understood. The new understanding applies not only to atherosclerosis but also to other illnesses characterized by abnormal interactions between blood and vessel walls, such as blood-clotting disorders, adult respiratory distress syndromes, and high blood pressure.

Structures, molecular mechanisms, and controls involved in various components of blood-blood vessel interactions have been identified. For example, the inner lining of blood vessels is a single layer of cells: the endothelium. Research has transformed our view of the endothelium from an apparently simple material with simple tasks to one capable of performing an impressive array of complex functions, among them the regulation of blood pressure, blood clotting, and the growth of new capillaries. The structure of the endothelium has been probed, as have the mechanisms by which materials cross it when moving from blood to tissue.

At the same time, the structure and functions of the blood components that interact with the endothelium have been investigated. These include platelets, essential to blood clotting; leukocytes, or white blood cells, which help to defend the body against infectious agents; and plasma lipoproteins, from which the cholesterol in atherosclerotic plaques is derived.

The impetus for research into these interactions stems from new concepts and techniques. For instance, more factors affecting the very complex set of reactions involved in the formation and removal of blood clots have been found. Two hypotheses concerning the origins of atherosclerosis—both dealing with the deposition of fat, especially cholesterol, upon inner arterial layers—have stimulated a wide range of research. New approaches to slowing the onset of atherosclerosis, ones coupling modified diets with medication, are being pursued. And new instrumentation for sorting and isolating cells now makes it possible to obtain sufficient numbers of individual blood cells for study. Finally, receptors and channels, the modes by which materials pass in and out of cells and cellular organelles, can be examined using advanced techniques of cellular and molecular biology.

Important clinical advances likely to emerge from this work include improved prosthetic devices, such as heart valves, vascular replacements, dialyzing membranes, and artificial organs, and fresh insights into the prevention and treatment of atherosclerosis.

Parasitism

Research on infections caused by parasites is driven by interlocking humanitarian and scientific motives. Parasitic diseases such as malaria and schistosomiasis affect more than a billion people globally. In the United States, the parasite, *Giardia lamblia,* is a common cause of epidemic diarrhea. Immigration, increased international travel, and the stationing of U.S. military and civilian personnel in countries where parasitic diseases are common are increasing the incidence of these diseases in Americans.

Confounding an effective attack on parasitic diseases is the very complex life cycle of parasites, which makes them extremely difficult to control without harming the host. In addition, parasites have evolved novel mechanisms for eluding the usual immunological and other defenses. However, these same traits—adaptability and complex life cycles—make parasites attractive for the investigation of such basic biological events as cell growth and differentiation. Thus, work on parasitic diseases has led to advances in molecular biology, immunology, membrane and cellular biology, biochemistry, and pharmacology. In turn, parasitology has benefited from advances in these fields by exploiting new techniques such as monoclonal antibodies and the isolation and copying of specific genes.

A constant theme in this field is exploration of the unique traits of parasites. For example, the usual response to most infections is the appearance of antibodies that can react with the surface antigens of the infecting agent. However, one type of parasite can change its surface coating hundreds of times during an infection, so that the antibodies invariably attack the wrong antigen. Research on such antigen structures has influenced research on gene expression already and may be important to understanding how genes are regulated. Continuing studies of parasitic evasions of immunological defenses may clarify the nature of such defenses, or the reasons for their absence, in other diseases. Exploring how antigens and other substances traverse parasitic and cellular membranes will enhance our understanding of membrane biology.

Finally, a better understanding of the basic biology of parasitic diseases should yield ways to combat them more effectively. Thus, studying the genes that modulate the transformation of parasites through their different stages may reveal new ways to interrupt their life cycles. Already, the use of monoclonal antibodies is leading to the development of greatly improved diagnostic reagents. And there are new methods to generate antigens for use in vaccines against those diseases.

Chemical and Process Engineering for Biotechnology

The phenomenal progress in molecular biology, genetics, and biochemistry in the last 20 years now makes it possible to have cells manufacture products ranging from simple molecules to complex proteins. The need today is for fundamental engineering knowledge—and people—to translate that capacity into commercial processes. That translation faces several difficulties:

- living organisms can mutate or change genetically, affecting process operations;
- biological processes must be completely aseptic; and
- these processes usually occur in very dilute, aqueous solutions, so the products have to be separated from large volumes of water; such separations are complicated by the fact that the products are often fragile, hard to purify, and structurally complex.

Surmounting these problems entails the design of suitable bioreactors for the large-scale culturing of plant and animal cells, the separation and purification of reaction mixtures in order to obtain products of sufficient purity at competitive costs, and the improvement of bioprocess instrumentation and control. Each area requires blending scientific with engineering knowledge. Bioreactor research, for example, involves the merger of such biological sciences as molecular and cellular biology, microbiology, and cell physiology with engineering skills, chemical kinetics, thermodynamics, fluid dynamics, heat and mass transport, and precise process control. Progress in separation and purification sciences necessitates in part adapting to large-scale processes such powerful techniques of the research laboratory as electrophoretic and affinity separations. The control of bioprocesses poses special demands, such as the on-line monitoring of complex products for which no sensors are available yet. Solutions to these problems may require the use of enzymes, monoclonal antibodies, and living cells as components of electrochemical and optical detectors.

Potential opportunities for applying biochemical technologies are diverse and provocative. In the area of human and animal health care, for example, a new family of products based on genetically engineered proteins may emerge that can detect quickly and accurately viral and bacterial diseases, susceptibility to autoimmune diseases, genetic defects, and neoplasms. Other proteins, such as those inhibiting the growth of tumors or those that dissolve blood clots, are being tested. In agriculture, the new technologies may yield fungicides and herbicides that are highly potent, specific, and environmentally safe. Other prospects lie in environmental protection, where biochemical engineering may provide methods of destroying or removing toxic products, and in the use of natural resources, such as the improved recovery of metals from low-grade ores.

Advanced Polymeric Composites

Bone, wood, and clam shells are natural composites: their structures have properties matched to specific purposes. A growing array of manmade polymeric composites similarly matches properties to use, most commonly to provide materials that, on a per-weight basis, are stronger and stiffer than the best structural metals. Such advanced composites are being used already in the manufacture of aircraft and sporting equipment, and are on the verge of major applications in automobiles, heavy equipment, robotics, and other areas.

The rapid technological development of advanced composites in the last decade has outpaced the underlying science. For instance, the understanding of the relationships between structure and properties is still primitive, as is the knowledge of why and how composite structures fail. A science of the design and processing of polymeric composites, embodying extensive computer-based modeling in design, engineering, and manufacture, needs to be developed. Further, the toxicity of components, their long-term environmental effects, and their reuse need to be studied and this knowledge applied to the development of new composites.

Advances on these and related fronts will amplify the already substantial use of advanced composites. For example, the wider use of composites in the automobile industry will depend on attaining an acceptable balance between processing speed and product quality, a useful technology for joining and repairing composites, and their long-term dimensional stability. The resultant benefits may be considerable. The costs of tooling for composites are much lower than for steel and allow for greater manufacturing flexibility, quicker design turnover, and less capital investment. Composites also are less likely to corrode than metals; lower vehicle weights will save fuel.

Supercomputer Architectures

Supercomputers, able to respond to about 100 million instructions per second, will soon be capable of executing 1,000 million instructions per second, rising to 20,000 million in the next decade. Such extreme speeds derive from the rapidly developing technology for raising the densities—and hence the speed—of integrated circuit chips; they also derive from new architectures which often embody, in a limited way, parallel or concurrent computations.

While further improvements in the underlying componentry are vital, attaining even faster speeds in the future will depend on implementing new computer architectures—specifically, the effective use of large-scale parallelism. This entails the development of computers that can execute many hundreds, thousands, or tens of thousands of instructions simultaneously, and of software that can orchestrate these simultaneous streams of computation effectively.

Faster computers, accompanied by refinements in software, will expand dramatically the applications of computers to ever more complex scientific and technological problems. To illustrate, computer simulation will affect aircraft design, the development of new pharmaceuticals, the design of energy storage systems and industrial products, and the testing of new generations of integrated circuit chips. In science, faster computers will be applied to simulating intricate phenomena lying beyond observation and experimentation. Examples include the path taken by an electron traversing a neutron star; a chemical reaction under extreme temperatures and pressures; the forces that give protons and neutrons their structure; the optimal conditions for a fusion reactor; the neural pattern triggered when, say, a finger touches an object; and climate, weather, and other atmospheric phenomena, such as tornadoes and wind shears. Finally, faster computers are vital to national security goals, weapons design, and to assessing phenomena such as "nuclear winter."

Information Technology in Precollege Education

The cognitive sciences—combining cognitive psychology, linguistics, philosophy, and biology—examine how humans process information. Artificial intelligence reflects a concern with how computers process information and their emulation of intelligent action and human perception. Recent progress in the cognitive sciences and in artificial intelligence, combined with more powerful, versatile, and accessible computers, provides a basis for new technologies to improve education. Problem-solving, hearing, and the organization of semantic memory are all areas to which the cognitive sciences have contributed substantially. Similarly, expert systems have provided both an original method for organizing the knowledge of a human expert and a window into the nature of human knowledge, skilled problem-solving, and reasoning.

These recent advances have occurred through a linking of the cognitive sciences, artificial intelligence, and educational research. This progress offers a major opportunity to create learning systems that can help students to acquire the knowledge and cognitive skills necessary for effective work and citizenship. Experimental learning systems such as DEBUGGY, an expert system for diagnosing a student's procedural errors in subtraction, are being tested already.

Analogous efforts certainly will not solve all—or even most—of the problems of education. However, they will provide a coherent and scientific basis for designing instructional systems and for training teachers and restructuring curricula. They also may create valuable new resources in the form of model electronic learning environments, while attracting a new cadre of professionals to education and to educational research.

Opportunities in Physics

Discoveries in physics have influenced virtually all of the sciences and have spawned industries. Observations of electrical and magnetic phenomena, starting in the eighteenth century, led to concepts that, in the nineteenth century, spurred a crescendo of experimental and theoretical knowledge of electromagnetism. This rich body of knowledge is the basis of electric power, telephony, radio, radar, and television. The emergence of quantum mechanics in the twentieth century underlies much of physics and chemistry and is the foundation of such technological discoveries as transistors, lasers, and solar cells.

Fundamental advances in physics continue to enrich all of science, and virtually all technologies. Cosmology and astrophysics are intertwined with the subnuclear physics of elementary particles. New discoveries in quantum mechanics have changed our knowledge of atoms and molecules dramatically and have revolutionized our understanding of solids. Semiconductors are not only the results of these discoveries, in the invention of the transistor and the solar cell, but also have led, in turn, to striking progress in many technologies.

Two contemporary examples convey the influences of physics. One is the deliberate structural design of materials through the arrangement of atoms in one or two dimensions. The resulting materials have remarkable properties quite different from those of natural materials, thus presenting scientific puzzles and technological opportunities. The physics of these layered materials is fundamentally interesting and their properties are technologically important to the computer and energy industries.

Another example is the contribution of physicists to biological problems, including recent work on transmembrane signaling—the transmission of information in brain, nerve, and muscle tissue. The molecular basis of such signaling is now accessible, and the perspectives of physicists joining with those of biologists are expanding upon a vast array of research questions, such as how nerves conduct information and execute commands.

These few examples illustrate continuing traits of physics: enormous diversity, the search for fundamental laws, strong connections to other sciences, and technological and industrial applications.

Solar-Terrestrial Plasma Physics

Plasma physics studies the interactions of charged particles with each other and with electrical and magnetic fields. Its research areas comprise, in addition to the effort to attain fusion power, the interactions of the sun and the earth: the chain of physical processes that starts with the generation of the sun's magnetic field in the solar interior and links it to activity at the sun's surface and, ultimately, to the earth's ionosphere and atmosphere.

Fundamental questions continue to drive solar-terrestrial physics. Why does the appearance of sunspots on the sun presage magnetic storms and auroras? What roles do magnetic fields play in stars and galaxies? Rapid progress toward answering such questions has occurred in the last decade, made possible by the increasing precision of measurements, numerical modeling, and the further development of theories applied to solar-terrestrial plasma problems.

Plasma phenomena in the solar system are mirrored in other stars, in the neighborhood of neutron stars and black holes, and in galaxies. The sun and the solar system have become, therefore, a laboratory in which astrophysical plasma processes can be studied *in situ* and with a precision attainable nowhere else. As a result, space and astrophysical plasma physicists have begun to work closely together and a new and broad research field is developing.

The power of the solar-terrestrial system as an astrophysical laboratory will be enhanced by the proposed multispacecraft International Solar-Terrestrial Physics Program. Together with the Solar Optical Telescope, this program can be expected to provide the fundamental underpinning of solar-terrestrial plasma research for the next 10 years.

The overall research goal is to synthesize growing knowledge to create a unified and quantitative model of events affecting the sun and the earth: solar wind, or the plasma connection between the sun and the earth; sources of coronal heating and solar flares; and links between solar activity and magnetic storms on earth.

PART II

An Outline of Selected Issues

This second part of the 1985 Outlook highlights several issues, abstracted either from the fields discussed in Part I or from reports and discussions within the National Research Council or the Committee on Science, Engineering, and Public Policy. The intent is to articulate national concerns involving science and technology. Within that broad purpose, two caveats apply to the issues discussed: (1) they constitute a selected rather than a comprehensive listing, and (2) they are described in outline rather than in detail, to keep this report brief.

Within these limits, the issues are:

- *international competition in science and technology;*
- *scientific and engineering personnel;*
- *cooperative work across disciplines;*
- *research and transportation;*
- *facilities and instrumentation;*
- *issues in genetic engineering;*
- *issues in human biology;*

- *scientific communication, technology transfer, and national security;* and
- *global atmospheric effects of nuclear explosions.*

International Competition in Science and Technology

As emphasized most recently in the report by the President's Commission on Industrial Competitiveness, the nation's ability to compete in global markets depends on several interlocking elements, among them the ability to create, apply, and protect new technology; an adequate supply of productive capital; a well-educated and flexible work force; and increased policy emphasis on international trade.

These multiple elements represent difficult tasks for legislators and policymakers. This section concentrates on one aspect of them: improving the nation's competitive strength in science and technology. It does this by using three examples taken from Part I of this report, all of them economically important: supercomputers, biochemical engineering, and advanced polymeric composites.

Supercomputers

Rapidly developing microelectronic technology and computer architectures have created the bases for major advances in computational speeds. Such revolutionary changes are crucial to maintaining U.S. leadership in many scientific and technological areas. However, they also will expose the U.S. computer industry to new international challenges as rapid fluctuations in product price and performance undermine the predictable customer preferences that have characterized the industry.

Given this context, it is essential that the United States look to the solidity of its technological position. That position needs to be measured continually against developments abroad and strengthened judiciously where weaknesses are found. Major supercomputer technology initiatives are under way in three agencies—the Defense Advanced Research Projects Agency (DARPA), the National Science Foundation (NSF), and the

Department of Energy (DOE). These activities need to be accelerated and coordinated carefully to ensure both a systematic exploration of significant design alternatives and a rapid translation of successful designs into commercial production. Also vital is the early involvement of the major user communities—especially the research universities—in the development of software and utilization expertise for the new machines.

Although DARPA has the leading role in funding major development projects in computing, the DOE role in the development of scientific supercomputing will be crucial also, especially in view of the latter's traditionally strong ties to the research community. In addition, the basic research programs of the NSF, which contribute to the conceptual and algorithmic bases for new styles of supercomputing, and the NSF computer access program, which will put the new machines into the hands of the broad scientific communities that must pioneer their use, are very important and need to be kept strong. Access to supercomputers is becoming indispensable to frontier research in a growing number of scientific and engineering fields, among them fusion research, quantum chemistry, particle physics, materials science, petroleum exploration, and process technology.

Early industrial participation in these developments is imperative. Means of increasing cooperation between the computer industry and university and national laboratory research groups should be explored vigorously by the federal agencies that fund major supercomputer development. To allow time for familiarization and formation of strong technology-transfer links, cooperation among the different sectors should be encouraged in the early stages of computer design. Administrative obstacles to research collaboration between companies and to the commercialization of experimental products need to be reexamined.

Extremely high computation rates often can be attained efficiently by tailoring electronic hardware to the requirements of particular computer-intensive applications. Such special design efforts have become a significant component of computer research that needs to be recognized explicitly and cultivated systematically. In this area, the breadth and many-sided ingenuity of the U.S. academic and commercial communities can be exploited to gain competitive advantage. To do this, high-quality design tools and fabrication systems need to be widely available. A component of the DARPA Strategic Computing Program will address this issue, but a supplementary NSF program aimed at making the resulting design facilities available to the entire U.S. computer science community also may be appropriate.

Increased attempts by the United States to learn from foreign developments, especially in Japan, are prudent in view of Japanese strength in certain lines of integrated circuit fabrication and current reports of rapidly growing capabilities in software. Much more systematic collection and translation of Japanese technical literature are called for.

Of course, there are other elements in the international competition in supercomputers that are not included in this brief discussion. These include:

- the appropriate role of government and industry in implementing the new computer architectures designed in the universities; for example, what would be the respective roles of government and industry in what is usually considered applied research and development?;
- problems arising from limited industrial access to supercomputers;
- assuring continuity for recent attempts by the federal government to increase access to supercomputers by academic scientists;
- financial and other incentives for U.S. companies to develop a new generation of supercomputers; and
- the level of software development needed to ensure optimal application of parallel architectures.

Biochemical Engineering

Several countries are trying to develop strong biochemical engineering industries. West Germany, Japan, and Great Britain have national institutes for biotechnology. Such investments are driven by the economic potential of biochemical engineering.

For example, it is estimated that global markets for biological products will run from $40 to $100 billion annually by the year 2000, or about 15 percent of the total annual market for chemicals.

The United States has a strong capacity for leadership in biochemical engineering, owing largely to the basic research conducted in American laboratories. Achieving that leadership requires a wider knowledge base than is now available, greater numbers of trained personnel, support for pilot studies of biochemical engineering processes, and working connections between basic biological research and engineering practice.

The knowledge needed has been summarized in Part I. Engineering personnel needs can be expressed as a shortage of both competent biochemical engineers and the faculty to train them. These personnel problems are worsening as biochemical engineering companies absorb both faculty members and recent graduates who have research and teaching talents. A second difficulty derives from the fact that many biotechnology companies tend to be small and oriented toward research and development, so that they do not have a sufficient variety of large-volume products to support the development of new pilot processes and large-scale production facilities. Further, the government, not currently a major buyer of biochemical engineering products, may see no reason for supporting pilot studies. The result may be a lack of both corporate resources and governmental rationale to initiate new production processes.

Overall, an issue for congressional consideration is strengthening the links between life scientists and biochemical engineers. Mechanisms might include:

- support for cooperative cross-disciplinary research;
- institutional grants to train graduate students;
- funds to enable academic units to purchase the equipment essential for contemporary research in biochemical engineering; and
- incentives for quality faculty to dedicate their careers to launching innovative university instructional and research programs.

Advanced Polymeric Composites

While the United States has a sound position in advanced polymeric composites, vigorous programs also are proceeding in Japan and West Germany. The United States is strong in the chemistry of these materials, in their materials engineering, and in their application; Japan dominates in many aspects of carbon fiber technology.

As with computers and biochemical engineering, the best response of the United States is not necessarily to mimic international competitors. Rather, the effective transfer of information among basic, applied, and developmental activities is needed, as are mechanisms to enable different disciplines to work cooperatively on materials problems. Such disciplines include chemistry, physics, mechanical and chemical engineering, materials science, computer science, and toxicology. Only about 30 universities have research programs in advanced composites; of these, only two have multidisciplinary groups in the area. There are only 40 full-time equivalent faculty members in this field nationally.

In terms of national policy, a major issue is the creation of several research centers devoted to basic research on advanced composites. The goals of such centers could be to:

- carry out high-quality scientific and engineering research;
- perform toxicologic assessments;
- provide scientists and engineers trained in specific disciplines for research on advanced composites; and
- infuse engineering curricula with new knowledge.

Issues for the Congress

These three fields—computers, biochemical engineering, and advanced composites—illustrate both special needs and general guidelines for maintaining their strengths. The general lessons for effective progress, which are applicable to other fields, include the need for:

- ***complementary competence both in the basic science and in the developmental engineering, including personnel trained in both the fundamental science and the engineering principles underlying new technologies; and***

- *mechanisms to link different disciplines with each other, universities with industry, and basic scientists with technologists.*

An additional issue for the Congress to consider is:

- *the extent to which the expansion of technological programs for defense is creating shortages of trained personnel in areas critical to our international competitiveness.*

Scientific and Engineering Personnel

Only a few issues are discussed under this broad topic. These issues include the real difficulties of a young investigator trying to begin a career in research; the paucity of clinician-researchers; possible shortages of trained research personnel some five years from now; and the role of foreign nationals in U.S. advanced education in science and engineering.

Starting a Research Career

There is, typically, a cyclical pattern to surpluses and shortages of trained research personnel relative to job opportunities. The system tends to adjust to small oscillations; on occasion, the swings become quite large and require national attention. Thus, we now face severe shortages of computer science and engineering faculties as a result of insufficient numbers of doctorates in these fields and the large competition from industry.

In contrast, upon completing their training in biomedical research, many young people cannot find suitable openings and support to continue their research careers. Specifically, the concern is with research trainees in their mid-20's to mid-30's; that is, those who are doing much of the experimental work in fast-moving research fields, such as those described in Part I—oncogenes, atherosclerosis, and parasitology. Similar difficulties were seen in physics in the early 1970's and in mathematics in the late 1970's. The NSF postdoctoral program in mathematics, instituted to prevent the loss of a generation of gifted young mathematicians, may be applicable to other fields of science.

Several consequences follow. Promising students may turn away from biomedical research in favor of more secure and remunerative careers. Some of the best academic departments admit and train far fewer individuals than their pool of qualified applicants, faculties, and facilities permits. The overall impact—as in other fields of science—eventually may be an insufficient flow of young people into research careers and slower progress in exploiting research advances.

The Institute of Medicine's Committee on National Needs for Biomedical and Behavioral Research Personnel observed that this problem cannot be solved solely at the training level and that it is addressed more effectively in terms of funds available to support faculty members and their research programs.

The cost of equipment to set up a new investigator in many branches of science and engineering now runs into hundreds of thousands of dollars. These funds must come from institutional resources. This precludes many universities from making appointments. Even those research universities with the greatest financial resources are finding it difficult to meet these costs. The result is a pattern of shifting away from bringing young investigators into the system in favor of attracting established investigators who are better able to bring external resources with them.

Clinicians in Research

A related concern is the declining number of clinicians entering research. Yet, clinician-researchers are indispensable for progress in areas such as the biology of atherosclerosis, discussed earlier. Current clinical training programs in universities offer both inadequate salaries to trainees and uncertainty of continued support. The fact that fewer clinicians are entering research undermines the transfer of basic research to clinical practice and lessens the contributions of physicians in directing research into the proper channels for understanding and managing human diseases.

Possible Shortages

Beyond these immediate problems affecting the sufficiency of research personnel, several more may be in the offing. Student enrollments reflect job opportunities. Thus, the numbers of bachelors' degrees awarded in computer sciences rose from 5,600 in 1976 to 15,000 in 1982, while in engineering the figures were 45,000 and 74,000, respectively. In contrast, the numbers of recipients of bachelors' degrees in mathematics dropped from 15,800 to 10,900 between 1976 and 1983; the corresponding figures in the biological sciences were 54,000 and 43,000, respectively. The diminishing pool of students from which candidates for doctorates and postdoctoral work in various fields will be drawn five years from now could lead to shortages of trained personnel for universities and industry in the 1990's.

Doctorates for Non-U.S. Citizens

Five percent more doctorates in science and engineering were awarded in the United States in 1983 than in 1978. Virtually all of this increase is accounted for by degrees given to non-U.S. citizens on temporary visas. Overall, about 20 percent of all doctoral degrees in science and engineering in 1983 were earned by those holding temporary visas. In the same year, more engineering doctorates were awarded to foreign citizens than to U.S. citizens; 38 percent of the doctoral degrees in mathematics and 35 percent in the agricultural sciences went to foreign students.

Overall, the proportion of master's and doctoral degrees awarded to foreign students relative to American students has increased substantially in engineering, but has plateaued or fallen slightly in the physical and biological sciences. In actual numbers, graduate enrollments in engineering increased from 36,000 in 1976 to 53,000 in 1983. Of these, there were 24,000 and 31,000 U.S. citizens, respectively. That reflects a 30 percent increase, compared to an 80 percent increase in foreign graduate students in engineering.

Much of the increase in numbers of foreign graduate students enrolled in science and engineering is a direct consequence of the normalization of relations with the People's Republic of China in 1979. Improvement of higher education in the People's Republic was adopted at the Fourth National People's Congress held in January 1975 as a part of Premier Chou En-Lai's doctrine of "four modernizations." In 1984, China (including Taiwan) led all other foreign countries in the number of doctorates awarded to non-U.S. citizens in engineering and science, with the exception of the social sciences. The number of doctorates awarded to Chinese nationals in electrical engineering in that year is particularly impressive.

How good are these foreign students? One measure is their scores on Graduate Record Examinations (GRE). On average, foreign students enter U.S. doctoral institutions with quantitative skills, as measured by the GRE, exceeding those of U.S. students. The differential is smallest for students in the mathematical and physical sciences and greatest for students in the social sciences. The higher performance of foreign students on the GRE quantitative examination may reflect the higher selectivity in the application and admission of foreign graduate students compared with U.S. students. Understandably, foreign students, for whom English is a second language, perform less well than do U.S. students on the GRE verbal examination.

There are benefits and possible costs to this major participation by foreigners in U.S. graduate education in science and engineering. The benefits to the United States are substantial; they include the exposure of a new generation of foreign scientists and engineers to American society and culture, the opportunity for American faculties and students to gain foreign perspectives on current research, and the improvement of international scientific and engineering communication. In some instances, the presence of foreign students has made up for low enrollments of American students and faculty shortages, and helped to meet industrial needs. Thus, a substantial proportion—56 percent in 1983—of foreign doctoral recipients are remaining in the United States, in academia, or industry.

There are also some possible costs. Foreign students who do not return home after being educated in the United States are a "brain drain," particularly for less developed countries. While the opportunity to remain in the United States usually presents greater opportunities for research, the home country is denied the benefit of science and technology transfer for development.

Further, a higher proportion of graduate teaching and research assistants for whom English is a second language reduces the effectiveness of teaching at American universities and may even deter promising American students from taking advanced degrees. The greater availability of foreign students, often with financial support in hand, may decrease the incentive to recruit promising American students. Finally, and more subtly, foreign students and faculty members often have a theoretical rather than an experimental bent, a difference that may affect both the directions of future research and the efficacy of programs intended to accelerate the use of knowledge.

Overall, it is in the U.S. interest to serve as "schoolhouse to the world," to provide graduate education in science and engineering to the brightest students, no matter where they come from. The real issue is *not* the number of foreign students training in U.S. graduate schools, but rather the reduced proportion of U.S. students taking advanced training, especially in engineering. Incentives are needed to make advanced training in several fields of science and engineering more attractive to U.S. students.

Issues for the Congress

Even this brief summary of several personnel issues in national science and technology makes it clear that public policy is only one factor in dealing with them. Others include market forces of supply and demand, economic cycles, individual perceptions of promising careers, and the policies of foreign governments with regard to their brightest students. Further, there are other issues of comparable importance, such as providing the fullest opportunities for women and minorities to contribute to the health and vigor of the research enterprise. Within these limits, there are a number of issues in this area for the Congress to consider:

- *federal programs and policies that would help to minimize the impact of and reduce the cyclical fluctuations in the mismatch between the supply and demand for young investigators;*
- *additional programs, complementing the Presidential Young Investigators Awards, to help young researchers begin their careers; and,*

- *in its review of immigration legislation, policies regarding the ability of promising foreign students to study in American schools, of U.S. industries and universities to employ them, and the impact of such policies upon U.S. institutions, corporations, and other nations.*

Cooperative Work Across Disciplines

Consideration of new funding modes, research structures, and agency organization may become a major legislative issue, driven by a growing need for multidisciplinary research. Thus, many of the fields discussed in Part I of this Outlook benefit from different disciplines working on common problems. This situation is not new, but its breadth is and its importance may be. Multidisciplinary work now ranges from neuroscience, requiring the effective collaboration of biologists, anatomists, physicists, physicians, chemists, cognitive scientists, and computer scientists, to deliberately structured materials, requiring atomic and condensed-state physicists, materials scientists and engineers, chemists, toxicologists, and process designers.

Such cooperative work is clearly in the national interest, scientifically and technologically. To be effective, it demands continuing disciplinary strength and flexible mechanisms. Although research involving the collaboration of chemists, biologists, engineers, and others is common in industry, such collaboration across traditional disciplinary boundaries needs nurturing to function optimally in academia. Universities are, understandably, conservative in creating new organizational structures and in starting new programs. An issue certainly will be how to encourage the creation of such mechanisms without sapping disciplinary excellence. The NSF program for Engineering Research Centers is one example whose effectiveness will be tested. Undoubtedly, others will emerge.

The creation by NSF of a program on the chemistry of life processes responds to a particular need in multidisciplinary research: funding work not falling into the established programs of support agencies. Another example is biophysical research of the sort discussed in the section on "Opportunities in Physics" in Part I.

The Special Case of Agriculture

One aspect of cooperative work is moving the techniques and insights of one field into another. This is well illustrated by the efforts of the scientific community since the early 1970's to apply the spectacular advances in molecular biology and the techniques of genetic engineering to agricultural research. In time, such research will offer major improvements in crops and cropping practices. Promising research areas include: understanding the genetic information system governing plant growth and function; the genetic and biochemical systems controlling the virulence of plant pathogens; biological nitrogen fixation; biological controls of insect pests; and photosynthetic energy conversion and carbon metabolism.

Such research opportunities are constrained in several ways, such as insufficient numbers of scientists trained in both molecular biology and genetics as well as in agronomy and the basic plant sciences. Inadequate methods for stimulating cooperative research among these and related disciplines and for exchanging ideas are another obstacle.

The competitive grants program for agricultural research, established by Congress in 1977, is one important device for overcoming such constraints. However, it continues to be underfunded when compared to other programs supported by the Department of Agriculture. The latter are either intramural or operate under a formula funding structure. The current scale of competitive grants is too small to support a competitive molecular biology laboratory focused on agriculture.

In summary, Congress can usefully continue to consider mechanisms, such as the competitive grants program, that will reduce disciplinary barriers among the agricultural sciences and between them and the biological sciences. Such barriers are especially unfortunate when compared with the way in which technologies such as gene cloning and recombinant DNA are able to unify other biological disciplines. These barriers continue to retard what is still an inadequate national effort in the development of the molecular biology of plants.

The Administration has supported strengthening the basic research capabilities of the Department of Agriculture. However, further opportunities exist, and the Department is encouraged to continue to enhance its capabilities in molecular biology and related areas. Attention must also be paid to the process by which work supported by the Department and other federal agencies enters agricultural development and practice.

Issues for the Congress

The impetus toward multidisciplinary research, including the particular needs of agriculture, suggests several issues:

- *articulating an appropriate federal role for encouraging multidisciplinary research;*
- *providing incentives for universities to undertake multidisciplinary research, whether within their campuses, with other universities, or with industry; and*
- *encouraging the infusion of basic biology into agricultural research.*

Research and Transportation

A chronic issue for Congress is maintaining and improving the national transportation system. There are difficulties in bringing the benefits of sound research and effective technological improvements to this vast system. It is neither easily nor quickly nor cheaply changed. Further, transportation systems must be considered both in terms of their components—such as changes in the speed limit on highways or in the vehicles using them—and also as a network of highly interdependent parts; for example, rail and marine systems or highways and air transport. These systems are also affected by issues involving the environment, the economy, and public safety. Finally, a host of uncertainties—regulatory and economic policies, antitrust, and so on—often tends to depress private-sector spending on research and development in transportation.

Nevertheless, various transportation modes have benefited from the introduction of new technologies. For example, improved fuel economy, better safety, and lower pollution levels have been achieved for cars and other vehicles. Improved container technology has helped to integrate freight transportation.

Transportation systems undoubtedly can use to advantage much of the research done in other sectors. Thus, the development of advanced composites, discussed in Part I, as well as progress in combustion and heat transfer technology, in computer-aided design and manufacturing, and in optical scanning techniques have improved, and will continue to improve, transportation systems. Mechanisms for effective information and technological transfer across transportation modes will benefit the nation.

Highways

A comprehensive examination of research opportunities for these transportation modes, both individually and in terms of the interactions among them, is clearly needed. The advantages of such research to the nation are likely to be considerable. That can be illustrated by examining, briefly, research opportunities for just one facet of transportation: the highway system.

Over the years, some $1 trillion has been invested in about 4 million miles of highway, about a quarter of which represents federal investments. The costs of repairing this road network have increased, but the chief source of revenue that finances these repairs—the tax on motor fuels—has been increased only once in real terms in the last two decades. With increasing efficiencies in fuel economy, the greater use of alcohol and other fuels usually exempted from tax, and other factors, future revenues are likely to decrease.

However, the point is not solely one of additional construction funding; indeed, Congress voted $58 billion in federal aid for highways in the Surface Transportation Assistance Act of 1982. The point is the emplacement of a serious and concerted research program to find better ways to build, maintain, and operate the highway system. The gap between costs and revenues can be closed in part by research to identify more durable highway materials than those used now and the factors that lead to their deterioration. About $70 to $75 million is spent annually on highway research, but most of it is parceled out in "problem-specific" contracts of $30,000 to $300,000, resulting in research applications that are highly local and often not applicable to the generic problems of highways. Responsibility for conducting such research tends to be distributed, and therefore diffused, among private and governmental groups at the local, state, and national levels. In addition, the research tends to be parochial: virtually no work is done on asphaltic materials, even though almost all roads and streets are surfaced or resurfaced with such materials.

Issues for the Congress

Summarizing a recent congressional review of the nation's transportation research activities, then Representative (now Senator) Albert Gore, Jr., observed that "considering both the importance of the nation's highways to commerce, industry, and recreational activities and the staggering estimates of repair and replacement costs—placed by some at a major part of the total cost of trillions of dollars—the need for and importance of a well directed and targeted research effort becomes clear." Representatives of industry have voiced similar sentiments, pointing to the large dollar paybacks expected from accelerated research in highway and bridge materials and construction. A recent report of the Transportation Research Board of the National Research Council supports these assertions and estimates that an annual investment of $30 million over 5 years in an interdisciplinary research program would translate into a saving of about $600 million each year from improvements in highway performance.

Among the immediate issues, then, are:

- *the creation of a research program on highway transportation that is coherent, durable, adequately funded, and of a quality commensurate with the level of national investment; and*
- *the need for systems studies to attain optimal utilization of the various components of the transportation system—air, marine, motor vehicle, and rail.*

Facilities and Instrumentation

The inadequacy of facilities and instrumentation in research universities is recognized by the Administration and the Congress. As the Congress prepares to deal with particulars of this issue in the next several years, some comments on the research infrastructure may be useful. These comments focus on three aspects: large-scale facilities; the intensifying need for medium-scale instrumentation; and the problems raised by recent and partly successful efforts to bypass normal agency procedures, including scientific and technical reviews, used to ensure merit.

Large-Scale Facilities

Experimental science increasingly entails large financial commitments for the acquisition and maintenance of instrumentation and specialized facilities. To undertake frontier research, some disciplines now require facilities with capital costs of several hundred million to several billion dollars. Recent high-cost proposals envisage new facilities for materials research comprising synchrotron radiation and neutron-scattering facilities, earthquake engineering facilities, major new astronomy facilities, and the superconducting supercollider.

The immediate issue is the scale of costs; more subtly, it is the effect of such large funding commitments on less costly instrumentation and facilities or on other research programs. Specifically, the latter concern is that attention to large facilities may overshadow the need for smaller-scale instrumentation and limit grant support. In budget terms, the belief is that, while large facilities may be nontransferable additions to the budget, medium-scale instrumentation must compete for marginal increases in existing budgets.

With intensifying pressures for both large-scale and medium-scale instrumentation, Congress will face several questions in the next several years: What approaches are optimal in the decision-making required for initiating and funding large facilities? Is it in the interest of the United States to encourage international funding for major facilities? If so, for which research areas? And in which areas should the funding be exclusively national? How is funding sustained over time, despite yearly appropriations cycles? Are the typical funding patterns the most effective; for instance, support scattered among several agencies? Is the concern that "big" science impoverishes "small" science valid? That is, does support for large-scale facilities, such as sources of synchrotron radiation, reduce funding for laboratory-scale materials research or support for individual university researchers? If so, what remedies are applicable? How can the individual university researcher be assured of the funds, including travel costs, required to use major facilities productively?

Medium-Scale Instrumentation

New directions in a number of fields are increasing the need for medium-scale instrumentation. Medium-scale instrumentation refers to tools that, in costing from $100,000 to $1,000,000, fall between the major facilities discussed above and the relatively lower-cost instruments accessible to most researchers. Purchases of medium-scale instruments inevitably entail special budgetary procedures, in many cases involving pooling mechanisms such as institutional and regional facilities. Further, most federal grants for medium-scale instrumentation require significant cost-sharing by institutions. Realistically, however, institutional resources, even those of the major research universities, are unable to meet all of the requests from their faculties for cost-sharing.

Grants to buy instruments need to be accompanied by funds for operation and maintenance throughout their expected lifetimes. That is not always the case today. Funds to purchase instruments often do not include money to operate them or to hire and train technical personnel to help investigators use state-of-the-art technology. Sometimes, it is easier for academic institutions to obtain financing for new equipment than for operating and upgrading existing equipment, even though the latter may be only a year old. Therefore, it would seem prudent for granting agencies to consider policies for funding capital equipment that allow money for maintenance, operation, upgrading, and training or hiring technical personnel.

In any case, concern for the inadequacy of advanced instrumentation for academic research is likely to increase in the next few years. There is a significant gap in access to medium-scale, technologically advanced instrumentation between academic and industrial researchers working in the same area. While several federal agencies—such as the National Science Foundation, the Department of Defense, and the Department of Energy—have initiated instrumentation programs to bring state-of-the-art equipment into university laboratories, these programs tend to be small compared to the magnitude of the problem. Also, they must be stable over several years.

To give the issue greater concreteness, instances of four fields offering rich research opportunities but constrained by inadequate, medium-scale instrumentation are discussed below. These examples are: chemistry, materials science, neuroscience, and biochemical engineering.

Chemistry. These are propitious times in such chemical fields as reactivity, catalysis, and the chemistry of life processes. But favorable opportunities can be exploited fully only with highly sophisticated instrumentation to create molecular beams, to follow the energy and compositional changes of a catalyst during a reaction, to watch reactions occurring more rapidly than a beam of light can cross a strand of hair, and to provide a host of techniques for separating and then identifying incredibly minute amounts of complex molecules.

The difficulty is that, historically, funding for basic research in chemistry did not allow for access to such sophisticated equipment. The resultant lag in instrumentation in universities threatens the dynamism and opportunities for chemistry in the United States and, ultimately, the international competitiveness of several of our major industries.

Deliberately Structured Materials. As mentioned in Part I, this field has advanced rapidly in the last 10 years because of unprecedented innovations in atom-by-atom control of both the synthesis and the characterization of materials. The potential applications are immense; communications and computers are the most obvious.

It is impossible to do research in this field without medium-scale instruments: X-ray apparatus, electron microscopes, special laser systems, ultrahigh vacuum systems, and so forth. A specific example is molecular beam epitaxy, a technique for atom-by-atom layering of a crystalline surface, creating a sort of molecular pousse-cafe. The technique enables exacting controls of the electronic and optical properties of a semiconductor or a catalyst. The paucity of molecular beam epitaxy equipment in university laboratories impoverishes both academic research in this field and interactions between academic and industrial researchers.

Neuroscience. Neuroscience is one of the most active and dynamic fields in science, yet it also faces serious equipment obstacles. For example, there is a major need to assess the active human brain noninvasively by methods that include radiological and nonradiological imaging, electrical and magnetic mapping of brain activity, and techniques in which brain function and response are assayed by chemical measurements of blood and spinal fluid. Funds for improving and developing equipment are needed to accomplish this work. Further, advances in solid-state measuring devices now available commercially have not been applied to research as fully as possible because of insufficient money to modernize laboratory equipment.

Biochemical Engineering. As indicated earlier, this field requires critical research elements in order to translate basic biology into large-scale processes. Advanced research demands, among other things, fermentation equipment and tissue culture laboratories, the latter costing approximately $250,000 each. That amount of money is not available in standard research grants. As in the case of materials research, the effect is to limit both academic research in bioengineering and dialogue between university and industry.

Planning New Facilities

Decisions by federal agencies and Congress on large expenditures of public funds for scientific facilities are necessarily tempered by costs and policies. Within these constraints, scientific evaluations have provided objective assistance in decisions to initiate major scientific facilities and instruments and in their selection, siting, and operation. This process has served the country well.

However, in the past several years, there have been—and continue to be—efforts to circumvent these important elements of consultation and open competition. Attrition of accepted and successful mechanisms threatens to: (1) eliminate scientific justification as an element in spending large amounts of federal funds, (2) disenchant the many able scientists who voluntarily and carefully review proposals for federal funding, (3) discourage other institutions already frustrated by the limited amount of federal money available for research facilities, and (4) erode the process for judicious allocation of funds among facilities, institutions, programs, and projects.

In thinking about the funding of facilities, it is useful to keep in mind that there are several classes of facilities and that the review and approval processes for different classes of facilities are hardly homogeneous, differing from class to class and from agency to agency.

Granting that the divisions are arbitrary, one can distinguish at least four classes of research facilities:

(1) national facilities, intended to serve a national, often international, research community—for example, the Fermi National Accelerator Laboratory;

(2) university-based research facilities—a new chemistry building, for instance;

(3) regional research facilities, usually based at a university—for example, the Triangle Universities Nuclear Laboratory in Durham, North Carolina; and

(4) technology centers, tied to local and regional economies

and located at or affiliated with universities—for example, the Basic Industry Research Institute at Northwestern University.

By and large, the review procedures for class (1), national facilities, are well established and have worked satisfactorily. That is true for Fermilab and for recent planning, such as for the new synchrotron radiation facilities now under consideration.

Recent and successful efforts by universities to circumvent normal comprehensive review mechanisms for deciding on facilities have involved the other three classes. A committee of the National Science Board estimates that, through the direct appropriation process (rarely used in funding academic research facilities), 15 universities have succeeded in obtaining over $100 million for facilities without going through the usual processes of open competition and review of scientific merit—two important elements in decision-making.

Many groups and individuals—in and out of government—have decried these circumventions. The potential for vital damage to the U.S. research system and to the apolitical role of the universities has been well articulated. The point is the nature of the pressures prompting such actions and what can be done about them. Certainly, a major source of such actions lies in two intersecting trends: (1) the decline, for over a decade, of federal support for new facilities and the renovation of existing ones, and (2) the explosive growth of science and, with that, the corollary need for facilities—with ancillary operating funds and personnel—to provide the complex instrumentation required at the frontiers of research. As the president of Stanford University recently observed in *Science*: "The political spasms we are now seeing result from the struggles of the scientific venture to escape from the prison of its own undercapitalization. . . . We now find ourselves caught in a mismatch between the needs and expectations of scientific research and the willingness of the public sector to support it."

Issues for the Congress

Given the large deficit in the U.S. budget and other constraints, it is unrealistic to expect the government to provide all of the large

shortfalls in capital and operating funds for academic research facilities. In any case, it is rare—national facilities of the Fermilab type again excepted—that the federal government covers all of the costs of an academic facility. The issue is how can federal funds for facilities be leveraged most effectively and equitably? What responses will improve the overall health of the U.S. research system?

First, the universities must be imaginative in obtaining funding for facilities, especially in leveraging private-sector, state, and federal investments. Illustrative mechanisms might include:

- *the use of federal funds to pay the interest on money borrowed by universities;*
- *the use of local bonding authorities;*
- *sharing costs with other universities and research centers; and*
- *establishing different overhead rates for different sections of the university.*

Second, on a national scale, the contending pressures of rapid scientific and technological advances, a decade of reduced federal expenditures on facilities, and limited national resources present an extremely prickly set of policy issues. Among them are:

- *enhancing systems for comprehensive merit review which include judgments from the scientific community on the quality of science likely to be conducted at a facility to be built with federal funds;*
- *anticipating and planning for future needs for large scientific facilities;*
- *seeking optimal ways to leverage federal investments in such facilities;*
- *considering criteria to test whether international cost-sharing is appropriate for a large facility;*
- *ensuring that support for large new facilities is not offset by inadequate support for qualified individual investigators;*
- *reviewing current federal fiscal and accounting policies with respect to the capital and operating costs of facilities; and*
- *evaluating alternative mechanisms for funding instrumenta-*

tion, including capital cost recovery as a part of the operating budget of a federal grant or contract.

Issues in Genetic Engineering

In a short time, recombinant DNA and the related techniques of molecular biology have generated remarkable advances in basic biological research. These efforts are on the verge of yielding promising applications in clinical medicine, veterinary medicine, agriculture, energy, and pollution control.

Several federal agencies are assessing the impacts of applications that require the use of genetically altered organisms in environments less well controlled than those in research laboratories. For example, the Environmental Protection Agency is trying to develop its own capabilities for assessing potential ecological and health effects, for monitoring organisms and genomes in various environmental media, and for devising effective control technologies. These efforts are overseen by the Federal Interagency Recombinant DNA Committee. The Recombinant DNA Advisory Committee of the National Institutes of Health still exercises the main responsibility for supervision of research protocols.

The need to coordinate the policies of the federal agencies whose responsibilities encompass various applications was recognized in the recent publication for comment in the December 31, 1984, *Federal Register* of a "Proposal for a Coordinated Framework for Regulation of Biotechnology," prepared by the Cabinet Council Working Group in Biotechnology. This proposal describes the policies of the major regulatory agencies that review biotechnology research and products. It also provides a regulatory matrix, outlining the applicable laws, regulations, and guidelines.

The pattern established in the late 1970's—federal guidelines for research and specific risk assessment experiments for determining potential ecological hazards—remains useful. Altered microorganisms can be tested in the laboratory first and then in well-controlled field sites, before release into the general environment. Using such test environments, scientists can devise effective means of monitoring the fate of the organism and its specific genes.

Issues for the Congress

Two major related issues continue to be:

- *maintaining oversight of what have been remarkably effective mechanisms for monitoring the use of recombinant DNA and related technologies; and*
- *assuring full and balanced consideration of requirements for public safety and the needs and opportunities of an emerging industry.*

Issues in Human Biology

Many discussions of new biological techniques involving reproduction and human genes have been subject to misinterpretation. A description of a potential application sometimes devolves, incorrectly, into an assumption that the application actually is possible or indeed is on the verge of being put into practice. Also, it is difficult to emphasize sufficiently that the spectacular successes achieved in isolating and copying certain human genes leave unsolved the much more difficult tasks of inserting these genes into the right cell and the right DNA position, and then having them function properly. Further, the differences between the genetic content of germline and somatic cells—the first transmissible from parent to offspring and the second not—are often lost in the discussion. Yet, such differences are vast in terms of the technical difficulties of gaining access to and engineering genes, the possible risks, and the accompanying social and ethical considerations.

There are two areas in which new biological techniques do or may have roles. First, the fertilization of human eggs outside the body and their subsequent implantation is a technology that is in use but that suffers from a *de facto* ban in the United States on research to understand it better. Second, consideration needs to be given to potential techniques, typically classed as genetic engineering, involving human somatic and germline cells.

In Vitro Fertilization and Implantation of Human Eggs

Based on data reported informally at a recent international conference, this technology is in substantial use already as a treatment for infertility. There are now about 200 centers worldwide. Of these 200 centers, 50 are in the United States and about 70 have been in existence over one year. About 500 children have been born through these techniques. Many more embryo transfers have been carried out—about 7,500—but pregnancy does not always ensue. About 24,000 human eggs have been collected at these centers. This has been possible, in part, because hormone treatments allow women to produce more than one egg cell per ovulation.

Against this reality of a robust technology is the fact that there are substantial gaps in knowledge of the underlying science, especially of the embryo. For example, a fairly new technique is to remove a few cells from a very early embryo, one that is still a solid sphere of cells, freeze the embryo, and check the separated cells for chromosomal abnormalities. Assuming no problems, the embryo can then be implanted. Several children have been born through this technique, and three clinics in the United States are using it. Yet, we have little basic knowledge of the effects of freezing and thawing on the embryo or of the removal of a few cells.

Part of the problem is a *de facto* ban on federally funded fetal research. All such research must be approved by a board of the Public Health Service. But the board does not exist; no members have been appointed.

*Genetic Therapy of Human Cells**

Engineering of somatic cells—such as bone marrow cells—is similar to that of germline cells—eggs and sperm. In each case, a gene being transferred must (1) be put inside the appropriate cell, and (2) be positioned properly into the cell's DNA. Beyond these generalities, there are major and important differences.

---

* Useful background on this issue, especially on therapy involving somatic cells, is provided in the recent report by the congressional Office of Technology Assessment entitled *Human Gene Therapy*.

Somatic Cells. Genetic therapy with somatic cells is likely to begin with bone marrow cells. These cells are accessible, can be grown in the laboratory, and can be transferred successfully into a patient's bone marrow. Further, there are very serious genetic diseases in which the genetic defect is both relatively simple (for example, a single defect involving a single gene) and may be treatable through bone marrow cells. In contrast, genetic defects involving multiple genes, brain and nerve cells, or several different kinds of cells are much more difficult to treat and are not likely candidates for initial therapies.

A spectrum of techniques exists for transferring genes into a cell. Each has its limits, with the use of certain types of viruses being the most promising at the moment. However, no methods exist yet to deliver a donated gene with certainty to the targeted region of the recipient cell's DNA; nor are there methods to regulate precisely the expression of a properly inserted gene.

These and other factors in genetic therapy with human somatic cells are articulated in a recent statement, entitled "Points to Consider in the Design and Submission of Somatic-Cell Gene Therapy Protocols," by the National Institutes of Health Working Group on Human Gene Therapy.

Germline Cells. Several laboratories have been able to transfer genes into fertilized mouse eggs, which then develop *in utero* into living animals. The transferred genes are expressed in different tissues of these "transgenic" animals and are inherited by subsequent generations. Such experiments are providing new insights in developmental biology and tumorigenesis. These procedures are now being extended to farm animals.

In the case of humans, no experiments of this kind have been attempted nor, in contrast to gene transfer in somatic cells, is there consensus on the potential usefulness of germline therapy in human genetic disorders. Aside from the critical ethical questions raised by heritable modification of the human germline, there are also severe technical limitations. For example, it is not possible yet to target genes into their correct position in recipient cells; therefore, gene expression is unpredictable and possible deleterious effects of the random insertion of genes cannot be excluded.

Issues for the Congress

Overall, the immediate prospects for germline therapy are nonexistent, and the long-term prospects are highly problematical. The outlook for somatic therapy is brighter, but it still faces major technical difficulties.

The issues in the area of human biology include:

- *maintaining oversight through the National Institutes of Health of plans for human gene therapy; and*
- *encouraging programs to enlarge fundamental understanding of fetal biology.*

Scientific Communication, Technology Transfer, and National Security

The scientific and technological strengths of the United States depend in part on the rapid and free exchange of information. There is concern in the defense and intelligence communities, however, that this openness may be of military benefit to the Warsaw Pact countries. Part of that concern, focused on academic research supported by the government, was addressed in September 1982, in a report of the Committee on Science, Engineering, and Public Policy (COSEPUP), entitled *Scientific Communication and National Security* (also known as the Corson report, after its chairman, Dale R. Corson).

The Corson Panel limited itself to questions involving academic science, leaving unresolved the complementary issue of the communication and transfer of industrial science and technology. It drew two major conclusions: first, a national strategy of "security by secrecy" is flawed because there is no practical way to restrict international scientific communication without disrupting domestic scientific communication, which inevitably weakens American capabilities in military and civilian technologies; and, second, that a national strategy of "security by accomplishment"—i.e., one that emphasizes protecting the U.S. technological lead by aggressively promoting scientific and technical productivity—is a far better alternative. The panel also outlined "gray areas,"—categories of technologies that, by their nature, could not be either completely open or totally classified.

Much discussion followed upon the release of the report, both within and outside the government. But implementation has not followed and the problem has remained unresolved.

The Department of Defense

In the spring of 1984, the Department of Defense (DOD) proposed an alternative policy that, in lieu of "gray areas," returned to a basic "black and white" approach whereby DOD research contracts would stipulate whether a particular project was open or classified. Many in the scientific community welcomed this as a positive development, although most have reserved judgment until the new policy is formally adopted and implemented.

In view of the prevailing federal policy on classification (that information must be restricted whenever there is reasonable doubt about the need for its protection), the government may be inclined to adopt a more conservative approach when deciding whether to classify militarily sensitive research—for example, on microelectronics or composite materials—that heretofore has been completely or partially unrestricted. Moreover, most research universities have standing policies against classified research projects on campus, except during national emergencies. Since most universities do not maintain secure off-campus facilities, there is the possibility that these institutions might withdraw from certain types of research sponsored by the Department of Defense.

The Department of Commerce

While the Department of Defense has been developing policies for the control of information resulting from federally funded research, the Department of Commerce has been working on a revision of the Export Administration Regulations, particularly the portion dealing with the export of technical data. This effort has been undertaken in parallel with congressional debate on new statutory language for the Export Control Act. These two initiatives may have a far greater impact on international technology transfer than the new DOD policy. Whereas the DOD policy affects only work done directly under contract to the department, changes in the export regulations affect everything involved in or related to any movement of products, processes, data, or expertise outside the borders of the United States. The Department of Commerce has acquiesced recently to DOD's insistence that the latter has the rights of review and timely refusal of export licenses.

From the narrow standpoint of scientific communication, the proposed revision of the Export Control Act creates the possibility that individual researchers would be required to obtain validated export licenses each time they planned to give a lecture, participate in a symposium, or work in a laboratory where foreigners were to be present and where so-called "militarily critical technical data" were to be presented or discussed. Any of these activities would constitute an "export." A similar obligation would exist whenever the research deals with critical technical data and the researchers plan to travel, to publish abroad, or, using this logic, to publish in U.S. journals read by foreign scientists. The net effect would be the regulation of scientific communication by the Export Control Act.

The implications of these new restrictions on the private sector may be equally profound. For example, the communication of technical data to and from the foreign offices, subsidiaries, affiliates, and suppliers of integrated, transnational corporations is central to their global nature. If these data flows were restricted, corporate research, development, and sales efforts would be inhibited. Similarly, the ability of a company to compete in world markets is based, in part, on its capacity to deliver a product, service (and related knowledge), or technical data in timely and unrestricted fashion. Particularly for the so-called "dual use" technologies (those that have both military and commercial applications), the new regulations may restrict severely the freedom of companies to reveal technical data or specifications at sales meetings or, in some cases, even to market a product. Many in industry have expressed their concern about these proposed changes in the export regulations. They have urged officials at the Department of Commerce and other agencies to engage in broader consultation and public debate before implementing them.

Another aspect of the problem pertains to the multilateral control of technology transfer by and/or between free-world, industrialized countries. Some in private industry argue that, because the policies of other industrialized countries toward their own companies are generally more liberal, current U.S. national security export controls succeed only in damaging the ability of American companies to compete for their share of the market for high-technology international trade and there is no gain from the security standpoint. This brings into question the effectiveness of CoCom, the International Coordinating Committee on Multilateral Export Controls, which has evolved over the years in a largely *ad hoc*, incremental manner.

Issues for the Congress

The sometimes contending and equally legitimate aims of protecting the nation's security, enhancing the U.S. competitive position in international markets, and protecting freedom of scientific communication raise difficult and durable issues. As a result, COSEPUP has initiated a new study to address those issues, entitled The Impact of National Security Controls on International Technology Transfer. *The panel, to be chaired by Dr. Lew Allen, Director of the California Institute of Technology's Jet Propulsion Laboratory, and former Air Force Chief of Staff and Director of the National Security Agency, will complete its work within the time frame of the 99th Congress. Hopefully, this study will assist the Congress in its efforts to develope clear policies to:*

- *protect the interests of national security;*
- *promote freedom of exchange of basic scientific information; and*

- ***reconcile corporate needs to transmit technical data and sell products and processes internationally with the requirements of national security.***

Global Atmospheric Effects of Nuclear Explosions

Several studies of the global atmospheric effects of nuclear explosions have appeared, among the most recent being the December 1984 report of the National Research Council (NRC) entitled *The Effects on the Atmosphere of a Major Nuclear Exchange*. These reports agree that a major nuclear exchange could alter the atmosphere seriously, at least for the short term. The precise effects depend on such variables as the season, the sites and yields of detonations, the altitudes at which explosions occur, the amount and nature of the smoke produced, and the ways in which soot particles are scavenged and rain out, especially from the upper atmosphere.

Such uncertainties make it difficult to state firm conclusions. As the NRC report pointed out:

> All calculations of the atmospheric effects of a major nuclear war require quantitative assumptions about uncertain physical parameters. In many areas, wide ranges of values are scientifically credible, and the overall results depend materially on the values chosen. Some of these uncertainties may be reduced by further empirical or theoretical research, but others will be difficult to reduce.

Why do these studies? One answer is that strategic thinking—and world opinion—may be affected already by the prospects of a "nuclear winter" and other possible climatic changes. As Herbert Simon writing in *Science* and others have pointed out, the strategic implications of the nuclear winter hypothesis have not yet been examined to the depth required. Whatever the implications, it is imperative that they rest on adequate information and well-based estimates. As the NRC report stated:

> Long-term atmospheric consequences imply additional problems that are not easily mitigated by prior preparedness and that are not in harmony with any notion of rapid postwar restoration of social struc-

ture. They also create an entirely new threat to populations far removed from target areas, and suggest the possibility of additional major risks for any nation that itself initiates use of nuclear weapons, even if nuclear retaliation should somehow be limited.

Some have argued recently that the nuclear winter hypothesis has implications for strategic defense, arms control, first-strike effects, target selection, and techniques of battle management. Given that the nuclear winter hypothesis has been overlooked for decades, are there other consequences of nuclear detonations not yet considered?

The need then is to identify all possible consequences, to narrow uncertainties, and to obtain credible, quantitative, and reasonably accurate estimates of what might happen to the atmosphere in a nuclear exchange. That research has begun; it needs continuing support.

Issues for the Congress

The immediate issues concerning research on the prospects of a "nuclear winter" are straightforward:

- *adequate support must be provided to conduct the research program called for in recent reports;*
- *all research findings should be public, within the legitimate constraints of national security; and*
- *the scientific community should be engaged fully, not only in planning and conducting the research program but also in appraising its quality and implications.*

Final Comment

This Outlook has highlighted important progress in some selected fields of science and technology. It has defined a number of issues and opportunities that relate to this progress and to national goals. The Outlook records new contributions to several problems once seemingly insurmountable, from cancer to the global devastation wrought by parasitic disease. American science and technology continue to display strength and leadership.

These essential qualities will be sustained only if timely attention is given to the basic resources needed, from instrumentation and facilities to the research climate and training for young investigators. In sum, the course of the nation's research system, and the magnitude of its contributions to meeting national goals, continues to depend on the wisdom, support, and guidance of the federal government.

Research Briefings

for the Office of
Science and Technology Policy,
the National Science Foundation,
and Selected Federal Departments
and Agencies

NAS
NAE
IOM

Research Briefings

for the Office of Science and Technology Policy,
the National Science Foundation, and Selected Federal
Departments and Agencies

Committee on Science, Engineering,
and Public Policy

National Academy of Sciences

National Academy of Engineering

Institute of Medicine

NATIONAL ACADEMY PRESS
Washington, D.C. 1983

NOTICE: The National Academy of Sciences was established in 1863 by Act of Congress as a private, nonprofit, self-governing membership corporation for the furtherance of science and technology for the general welfare. The terms of its charter require the National Academy of Sciences to advise the federal government upon request within its fields of competence. Under this corporate charter, the National Academy of Engineering and the Institute of Medicine were established in 1964 and 1970, respectively.

The Committee on Science, Engineering, and Public Policy is a joint committee of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. It includes members of the councils of all three bodies.

This work was supported by the National Science Foundation under Contract PLN-8218330.

Available from

Committee on Science, Engineering, and Public Policy
National Academy of Sciences
2101 Constitution Avenue, NW
Washington, DC 20418

Printed in the United States of America

COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC POLICY

GEORGE M. LOW (Chairman), President, Rensselaer Polytechnic Institute
SOLOMON J. BUCHSBAUM, Executive Vice President, Customer Systems, Bell Telephone Laboratories, Inc.
EMILIO Q. DADDARIO, Wilkes, Artis, Hedrick and Lane, Attorneys at Law
ELWOOD V. JENSEN, Professor and Director, Ben May Laboratory for Cancer Research, University of Chicago
ALEXANDER LEAF, Chief of Medical Sciences, Massachusetts General Hospital, and Jackson Professor of Clinical Medicine, Harvard University Medical School
GARDNER LINDZEY, President and Director, Center for Advanced Study in the Behavioral Sciences
J. ROSS MACDONALD, William Rand Kenan, Jr., Professor of Physics, Department of Physics and Astronomy, University of North Carolina
JOHN L. McLUCAS, President, World Systems Division, Communications Satellite Corporation
ELIZABETH C. MILLER, WARF Professor of Oncology, McArdle Laboratory for Cancer Research, University of Wisconsin
GEORGE E. PALADE, Chairman and Professor, Section of Cell Biology, Yale University School of Medicine
JOSEPH M. PETTIT, President, Georgia Institute of Technology
LEON T. SILVER, Division of Geological and Planetary Sciences, California Institute of Technology
HERBERT A. SIMON, Professor of Computer Science and Psychology, Department of Psychology, Carnegie-Mellon University
I. M. SINGER, Professor, Mathematics Department, University of California, Berkeley
F. KARL WILLENBROCK, Cecil H. Green Professor of Engineering, School of Engineering and Applied Science, Southern Methodist University

Ex Officio
FRANK PRESS, President, National Academy of Sciences
COURTLAND D. PERKINS, President, National Academy of Engineering
FREDERICK C. ROBBINS, President, Institute of Medicine

COSEPUP Staff
ALLAN R. HOFFMAN, Executive Director
BARBARA DARR, Administrative Assistant

PREFACE

The seven research briefings contained in this document were prepared in response to a request from Dr. George A. Keyworth, Science Advisor to the President and Director of the White House Office of Science and Technology Policy (OSTP). The effort was directed by the Committee on Science, Engineering, and Public Policy (COSEPUP), a joint committee of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

Topics for the seven research briefings were selected by OSTP. For each topic a balanced panel of 11-13 experts was organized to develop the briefing. The specific charge to each panel was to critically assess its field and to identify those research areas within the field that were likely to return the highest scientific dividends as a result of incremental federal investments in FY 1984. It was also emphasized that these briefings were not to be construed as substitutes for the much more detailed surveys occasionally undertaken in major scientific fields (e.g., the recent report of the National Research Council's Astronomy Survey Committee entitled Astronomy and Astrophysics for the 1980's, Volume 1).

Through discussions with OSTP, the seven topics were defined as follows:

1. MATHEMATICS: Research covering the following fields of investigation: statistics, pure and applied mathematics, mathematical systems theory, numerical analysis, operations research, computational mathematics, and scientific computing.

2. ATMOSPHERIC SCIENCES: The study of the physical, chemical, and dynamic properties of the atmosphere and its interactions with the Earth, the oceans, and the planetary environment with a view to understanding and predicting the atmosphere's changes and behavior as manifested in weather, climate, air quality, and other characteristics relevant to human society, both as a result of natural processes and as influenced by human activities.

3. ASTRONOMY AND ASTROPHYSICS: Research with the objective to obtain information about astronomical bodies by remote sensing from the surface of the Earth, from the Earth's atmosphere, and from Earth orbit.

4. AGRICULTURAL RESEARCH: Research of greatest promise for increasing the productivity and efficiency of American agriculture, including:

- plant sciences targeted on developing more productive, resistant, tolerant, and energy-efficient crop plants;
- an objective assessment of the realistic expectations of genetic engineering for developing more productive, resistant, tolerant, and energy-efficient crop plants; and
- research on crops and cropping practices that are more resource conservative.

5. NEUROSCIENCE: Research directed toward understanding the molecular, cellular, and intercellular processes in the central nervous system (CNS) and the way in which those processes are integrated in CNS functional control systems, with emphasis on research relating CNS functions with behavior.

6. HUMAN HEALTH EFFECTS OF HAZARDOUS CHEMICAL EXPOSURES: Research on the responses of organisms to hazardous chemical exposures, including:

- the nature of the steps leading to damage to the organisms;
- the mechanisms and kinetics of metabolism of hazardous substances;
- the nature of protective responses and repair mechanisms; and
- the in vitro, animal bioassay, and epidemiological methods used to characterize hazardous exposures.

7. MATERIALS SCIENCE: Research concerned with reaching a clearer understanding of the complex relationships that exist among the atomistic structure, composition, and defects of materials and their behavior in an engineering environment. Specific areas of investigation include those concerned with surface characterizations, defect structure, electronic structure, catalysis, the theory of crystalline solids, and the properties of solids (e.g., electrical, magnetic, optical, thermal, and mechanical).

Each panel met once, for 2 or 3 days, to carry out its charge. Knowledgeable representatives of government and the private sector were invited to provide input to the panels. Rapporteurs, knowledgeable in the field, were present to summarize the discussions and prepare initial drafts of briefing papers. These papers were reviewed and revised by the panel members and served as the bases for the oral briefings presented to federal officials.

The seven one-hour briefings, presented by panel chairmen and, in some cases, 1 or 2 other panel members, were reviewed by COSEPUP in mid-October and presented to Dr. Keyworth and members of his staff between October 26 and November 18, 1982. The same briefings were subsequently presented to Dr. Edward Knapp, Director of the National Science Foundation, and other Foundation officials on three days in December. Briefings for other interested departments and agencies were held separately on the same dates.

None of this would have been possible without the financial support of the National Science Foundation and the cooperation, under difficult time constraints, of the panel members and staffs. We are indebted to both groups.

If judged useful to federal decision makers, the seven initial research briefings developed as an experiment in 1982 could serve as the basis for research briefings on other major fields of science in future years. Such briefings could supplement other inputs and become important new channels for communication between the federal government and the scientific community.

George M. Low, Chairman
Committee on Science, Engineering, and
Public Policy

CONTENTS

REPORT OF THE RESEARCH BRIEFING PANEL ON MATHEMATICS 1

REPORT OF THE RESEARCH BRIEFING PANEL ON ATMOSPHERIC SCIENCES 19

REPORT OF THE RESEARCH BRIEFING PANEL ON ASTRONOMY AND ASTROPHYSICS 31

REPORT OF THE BRIEFING PANEL ON AGRICULTURAL RESEARCH 57

REPORT OF THE RESEARCH BRIEFING PANEL ON NEUROSCIENCE 81

REPORT OF THE RESEARCH BRIEFING PANEL ON HUMAN HEALTH EFFECTS OF HAZARDOUS CHEMICAL EXPOSURES 99

REPORT OF THE RESEARCH BRIEFING PANEL ON MATERIALS SCIENCE 111

REPORT OF THE RESEARCH BRIEFING PANEL ON MATHEMATICS

Research Briefing Panel on Mathematics

William Browder (Chairman), Princeton University
Herman Chernoff, Massachusetts Institute of Technology
Hirsh Cohen, IBM Watson Research Center
Louis N. Howard, Florida State University
Arthur M. Jaffe, Harvard University
Peter D. Lax, New York University
Joel Lebowitz, Rutgers University
G. D. Mostow, Yale University
Michael Rabin, Harvard University
I. M. Singer, University of California, Berkeley
Shing-Tung Yau, Institute for Advanced Study

Guests

Donald Austin, U.S. Department of Energy
Stuart Brodsky, U.S. Department of Defense
Jagdish Chandra, U.S. Army Research Office
David Fox, Mathematical and Information Sciences, Bolling Air Force Base
Kenneth Hoffman, Massachusetts Institute of Technology
E. F. Infante, National Science Foundation

Staff

J. K. Goldhaber, Executive Secretary, Office of Mathematical Sciences
Allan R. Hoffman, Executive Director, Committee on Science,
    Engineering, and Public Policy

REPORT OF THE RESEARCH BRIEFING PANEL ON MATHEMATICS

PREFACE

In response to an invitation from OSTP to present a research briefing on the current state of mathematical research in the United States, a panel was convened on September 25 and 26, 1982. In addition, opinions were solicited from the chairmen of the top 27 research departments of mathematics in the United States.

This document is an account of the deliberations of the panel. Our charge was to identify special opportunities in the mathematical sciences. We could identify many promising areas ripe for development, but it is our belief that the most dramatic mathematical tendency in recent years is the drawing together of mathematics (often the most abstract science) with other sciences and the interplay between them. Some results have already been achieved; with encouragement, the interplay can be made increasingly fruitful.

A healthy mathematical enterprise must have effective means for nurturing new developments. These means involve (1) support of gifted young investigators who will choose from among the new directions and (2) flexibility for scientific leaders allowing them to develop and expose recent breakthroughs. Unfortunately, severe underfunding has limited the mathematical community's capacity to respond in these ways. For this reason, in addition to spelling out the current state of mathematical research, we have tried to specify the additional resources that will be needed to exploit new opportunities.

INTRODUCTION

There is a striking contrast between the importance of mathematical sciences in the United States and the perception of them. On the one hand, mathematics and its applications play an ever increasing role in science, technology, business, and everyday life in this country, and, on the other, mathematical research is almost completely unknown to and poorly understood by the general public and even the scientific public.

The reputation and achievements of the American mathematical community make the United States first in the world in mathematics. Yet at the same time support for mathematical research erodes at a steady rate, and the institutional infrastructure that supports the enterprise exhibits symptoms of decay. In a year in which two out of three of the quadrennial Field's medalists (the mathematical equivalent of the Nobel Prize) are American mathematicians, we find highly ranked departments of mathematics lacking enough research support to send their most productive people to professional meetings or to photocopy important documents. At a time when the development of research in mathematics is making unparalleled progress, when the influence of mathematics is pervasive in the other sciences, when American mathematicians lead the world in most areas, research support for graduate students and young Ph.D.s, the lifeblood of the enterprise, is insufficient to ensure the quality of future generations.

In Section 1, we describe some of the exciting contributions and important developments in mathematical research and the highly promising opportunities that arise from the recent opening of new bridges between mathematics and other sciences. In Section 2, we document the decay of the infrastructure and support of mathematical research, and indicate possible actions to prevent further decay of the mechanisms that have enabled U.S. mathematics to flourish. In Section 3, we recommend steps to be taken to rehabilitate the ailing infrastructure and to provide the flexibility needed for the exploitation of new opportunities.

SECTION 1 SOME RECENT DEVELOPMENTS IN MATHEMATICAL RESEARCH

Our exposition will be in two parts, the first giving examples of the pervasive influence of mathematics in other sciences and the second recounting some of the recent significant advances in theoretical mathematics. Some speculations on future possibilities will be interspersed. (For a somewhat fuller review, we refer the reader to the chapter "On Some Recent Developments in Mathematics" in Outlook for Science and Technology, The Next Five Years, NAS, W. Freeman & Co., San Francisco, 1982, pages 467-510.)

Influence and Applications of Mathematics

The research area of mathematics most used today in technology is numerical analysis and mathematical modeling. In the area of industrial design, for example, a given process must be described and understood in a mathematical way, and the details of the mathematical description will interact with the design process. Analysis and design become mathematically interdependent. For example, the design of the fuel efficient, weak shock transonic airfoil, currently flying on the Boeing 767, would not have been possible without the mathematical work of Garabedian, Cole, and Jameson. On the speculative side, a project to mathematically model the human circulatory system now under way might eventually have important medical consequences. Among them would be the possibility of indirect ways to measure the heart in a situation where the direct measurement of the heart itself is impractical. Computer-aided design (CAD) is now used in the design of artificial heart valves, using a mathematical model of the left side of the heart.

Mathematical design of efficient compression and turbine blades is a reality today, while the design of efficient combustion chambers is a subject of intense research.

In national defense, the replacement of experimentation by numerical modeling, made possible by advances in computers and dramatic improvements in mathematical algorithms, has resulted in great savings in the cost and improvement in the quality of design. This has been particularly significant in weapons-related research and development, where experimentation is costly, dangerous, and physically impossible in the early stages of a project.

In economics, mathematics is playing an ever increasing role, as witnessed by three recent Nobel prizes in mathematical economics.

In oil prospecting, mathematical results are used in a fundamental way in the separation of primary signals from multiple reflections. The modern theory of inverse scattering is becoming a basic tool in this area. Mathematical modeling is important in the study of efficient secondary oil recovery.

In electrical engineering, the mathematical work of Weiner has proved fundamental in several areas, and mathematical control theory plays an important role.

In medicine, great advances in diagnostic techniques (tomography--the CAT scanner; and NMR) are strongly related to mathematical research. In the latter, methods from singular integrals, complex function theory and Hilbert space were used. Statistics and statistical methods are crucial in epidemiology, drug testing, and many other areas, and mathematical modeling is an important tool in the development of new drugs. The list could be extended indefinitely with examples drawn from biology, chemistry, neuroscience, and other sciences.

There are many recent examples where mathematical research, driven by the inner dynamic of the subject without reference to practical problems, has been found to be of great significance in other areas. An outstanding illustration of this has been in the development of Gauge Field Theory in physics. Nobel Prize winner C. N. Yang wrote, "I found it amazing that gauge fields are exactly connections on fibre bundles, which the mathematicians developed without reference to the physical world." Algebraic geometry produced all self-dual solutions for the Yang-Mills equations. But the physical theory led also to important consequences in topology, as we will relate later.

Other new and important mathematical inputs into physics have been the introduction of abstract probability into statistical mechanics and material science with the notion of Gibbs state and the input of the theory of dynamical systems and ergodic theory into the study of turbulence. All these phenomena illustrate the drawing together of abstract and applied mathematics and their fruitful interaction.

As a result of the rise of the computer, the theory of computation has become an area of mathematical research. Solidly based on methods and fields in the mainstream of modern mathematics, such as probability, combinatorics, algebraic geometry, and number theory, it creates important tools for the practicing computer scientist. The main themes in this new area are the study of algorithms and of programming.

Efficient algorithms often have important practical importance. Notable examples are the Fast Fourier Transform with its application to signal processing and the recently developed randomized algorithms in number theory and finite fields with their application to error correcting codes and cryptography.

Developments in coding and cryptography provide dramatic examples of unexpected applications of "pure" mathematics to applied areas. Number theoretical work of A. Weil in 1948 was applied to coding theory some years ago. Last year a group of Soviet mathematicians showed how to use the latest work of Deligne, Rapoport, Ihara, and Langlands, in the most abstract areas of algebraic geometry, in the design of error-correcting codes of a theoretical efficiency heretofore deemed impossible.

In the field of robotics, the development of automatic industrial processes depends on successful mathematization or modeling of the processes involved. In many industrial areas, progress is in its infancy, and some of the most simple tasks seem the least likely to yield to automation. It is extremely difficult to design a robot arm with sensors that will enable it to avoid obstructions while picking up a target object, one of the most routine of human abilities. The parameters of this problem can be interpreted as a problem in algebraic geometry, and progress here may have some effect on the solution of other practical problems.

There are current proposals for initiatives in large-scale scientific computing that would have sizable components of mathematical research and important applications to applied mathematics. In pure mathematics, Thurston (one of this year's Field's Medalists) has made a surprising use of the computer as an experimental tool in his work on topology of 3-dimensions, although the solution of the famous four-color map problem a few years ago required the computer essentially in the proof.

Recent advances in computer technology and software are having a deep influence on the nature of work subjected to statistical analysis, on the methods of analysis, and on theoretical questions in statistics. The computer and space technologies provide vast amounts of high dimensional multivariate data which standard classical methods no longer fit, because the underlying assumptions of normality and linearity are no longer satisfied. Methods justified under those assumptions would lead to serious errors. Novel methods of pattern recognition and robust regression and new methods of graphical representation that allow the comprehension of these data are being developed. The interactive kinematic displays of Friedman and Tukey are examples.

New interactive statistical packages with properly built-in diagnostics will permit naive users to observe phenomena formerly accessible only to trained and ingenious statisticians.

Statistics has aided computer science in that simulations and experimental designs have been used to find good hardware configurations and software designs.

Progress in Theoretical Mathematics

By "theoretical mathematics" we mean research motivated by the inner dynamic of the subject rather than by the needs of research in other sciences. It is remarkable how often much of this research, seemingly irrelevant, turns out to have important practical impact. Who, for example, in the 1920s and 1930s would have predicted that the most abstract work in mathematical logic--recursive functions and "Turing machines"--would provide the philosophical framework for von Neumann's introduction of the stored program computer, in which the instructions to the machine can be manipulated and modified by the machine itself? The development led eventually to a multibillion dollar industry.

Progress has been spectacular in recent years and we cite some notable examples:

The work of Deligne, proving the famous "Weil conjectures" in number theory.

The classification of finite simple groups, as the end of a 20-year effort.

The work of Yau on the Calabi conjecture, with important applications to algebraic geometry.

The work of Thurston, showing how to employ methods of (mostly non-Euclidean) geometry to attack problems in the topology of 3-dimensions.

The work of Khachian on polynomial algorithms in linear programming.

The discovery of solitons and strange attractors.

The work of Connes on operator algebras.

Instead of continuing this list at length, we give in greater detail short accounts of some recent dramatic examples, starting with a startling advance in pure mathematics resulting from interaction with physics.

- Physicists have introduced gauge theories in 4-dimensions (space-time) as a unifying principle in field theory. The study of Yang-Mills equations of motion in this context led S. Donaldson to a remarkable description of certain 4-dimensional spaces. A little earlier, M. Freedman, using purely topological methods, had produced a powerful comprehensive theory of 4-dimensional manifolds. These results of Donaldson and Freedman have combined to give the following result in the topology of 4-dimensional space. In all other dimensions there is essentially one mode of doing calculus in a Euclidean space (R^n has a unique differential structure for $n \neq 4$), but an entirely different situation exists in dimension 4 (there are at least two different structures on R^4). This qualitative difference between dimension 4 and other dimensions is a startling development for topology, and it may also be the reflection of some deeply significant physical principles.
- The unifying role of group symmetry in geometry, so penetratingly expounded by Felix Klein in his 1872 Erlanger Program, has led to a century of progress. A worthy successor to the Erlanger Program seems to be Langlands's program to use infinite dimensional representations of Lie groups to illuminate number theory.

That the possible number fields of degree n are restricted in nature by the irreducible infinite dimensional representations of GL(n) was the visionary conjecture of R. P. Langlands. His far-reaching conjectures present tantalizing problems whose solution will lead us to a better understanding of representation theory, number theory, and algebraic geometry. Impressive progress has already been made, but very much more lies ahead.

Closely related to the Langlands program is the remarkable and mysterious connection between counting points in finite spaces and computing the topological invariants of continuous spaces. First propounded in the Weil conjecture, the connection is being made more accessible by the Goreski-MacPherson-Deligne homology theory. The whole thrust of these developments is to force the next generation of mathematicians to embrace heretofore widely separated areas of mathematics. The expected unifications are awesome.

- In analysis, the old problem of the regularity properties of the Cauchy integral (for Lipschitz curves) was recently solved by the work of Calderón, Coifman, McIntosh, and Meyer. Crucial to the solution of this problem were the techniques of Hardy spaces developed within the last decade, as well as recent methods for dealing with singular integrals with "rough" coefficients. It seems very likely that these ideas will be applicable to a host of important problems in partial differential equations, as is indicated by their role in recent advances in the solution of "Kato's conjecture" (dealing with square roots of Laplacians) and solutions of parabolic equations with minimal smoothness assumptions.

SECTION 2 THE MATHEMATICAL RESEARCH ENTERPRISE

In analyzing the state of mathematical research and its needs, we must keep in mind special features that distinguish mathematics from the other sciences. Among those features are these:

1. Mathematics is the most labor intensive of all sciences. Little equipment is involved, except for computers, which are heavily used in statistics and areas of applied mathematics and are an experimental tool for a few pure mathematicians.
2. The vast majority of research mathematicians are employed in universities as teachers. Industry and national laboratories support only a handful.
3. Very few federal agencies support research in mathematics. NSF supports 60 percent of all research in mathematics and almost 100 percent of pure mathematics; most of the rest is supported by DOD and DOE. This contrasts strongly with other disciplines.
4. The magnitude of the total research support of mathematics by the federal government in comparison with other fields is miniscule, less than $60 million annually.
5. Mathematics is "small science." Though collaboration among 2 or 3 researchers is not uncommon, large projects with many researchers devoted to specific goals are relatively rare. Mathematics thrives on the interaction of independent viewpoints and different approaches.

6. The health of the mathematical enterprise in the United States hinges on the strength and vitality of the departments in the leading research universities.

The following special factors have strongly influenced the pattern of decay in the support of mathematical research that we perceive.

a) It is now generally accepted that the impact of inflation is much greater in labor-intensive enterprises than in the general economy. The impact of declining resources and inflation has, therefore, been most severe in mathematics.

b) The universities as a whole are subject to the same effect, so that the resources available to them as the main supporters of mathematical research have dwindled proportionately.

c) There has been no organization of mathematicians expressing their discipline's support requirements for research. In sciences needing large instruments or projects to achieve their scientific goals, organized support has evolved and served effectively. But in the small-scale individualistic atmosphere of mathematics, no mechanism has evolved for calling attention to the alarming decline in funding.

d) Inflexibility is inevitable when few funding agencies support mathematical research; investigators not supported by NSF, for example, often have no place else to turn, unless their research has clear potential relation to the goals of a mission-oriented agency.

e) The number of top-ranking graduate students seems to be declining, and many of them are from abroad.

f) The strength of some of the leading departments of mathematics is being undermined by the lack of federal funding for research, a lack that the universities cannot replace, especially in states whose economies are suffering.

Yet the small scale of the mathematical enterprise would make it rather inexpensive to alleviate many of the serious shortcomings in the support configuration and to ensure the health and vitality of American mathematical research into the next century.

As the support of mathematical research by federal agencies has eroded over the last decade, the universities have, to some extent, taken up the support, as for example with postdoctoral research instructorships. The support has, however, become more and more difficult for the financially troubled universities to continue.

Support in the Mathematical Sciences

Currently, the United States ranks first in the quality of research in the mathematical sciences. But the vitality conceals a variety of problems that, if left unsolved, will inevitably lead to a substantial deterioration in the nation's mathematical sciences research enterprise. The same can be said substantially of other sciences, but it is our purpose here to document the special strains in mathematics that are reaching the crisis stage.

The consequences of the financial stringency are falling most heavily on the young mathematicians, graduate students, and recent recipients of the Ph.D. Little research support is available to graduate students, though many teaching assistantships are available, particularly at the large state schools. This means that mathematics graduate students, unlike those in other sciences, seldom have the opportunity to work full time in research. The lack of postdoctoral research appointments in mathematics creates a similar problem for the young Ph.D.s, who, in addition, are finding it increasingly difficult to obtain research grants. At a time of real opportunities, in the drawing together and mutually fruitful interaction of mathematics and applications, the young innovators who will be needed to exploit the opportunities are not receiving the nourishment they need for full development.

The chairman of a prestigious mathematics department writes (in a letter to the Panel): "Mathematical research has been flourishing in the past decade, but the institutional structure of mathematical research is in trouble. Recruitment of young talent for the future looks to be in even more serious trouble. The level of research support has been very low in terms of the percentage of active research people supported, and recent cuts in support have produced signs of a serious deterioration of morale especially among younger mathematicians."

Another chairman writes: "We are some one hundred in number, we are invariably ranked among the top twelve departments in the country, we continue to recruit good graduate students, and I claim with confidence that of the one hundred at least ninety are seriously engaged in research and scholarship. Yet, after two severe years, we are down from one-half to about one-third of the faculty on NSF grants. Moreover, we have sustained these severe losses without any sense of the prevalent quality of work having declined at all; on the contrary, several colleagues have lost grants in the very year when they have done their best work. . . . At the same time, universities are increasingly affected by lack of money. Here, for example, loss of NSF grants has reduced departmental income from overhead just when the university, which in any case had always counted on strong departments like ours to earn much of its research support outside, is quite unable to raise the level of state support. Also, of course, we find little endowment money coming in earmarked for mathematics."

Yet another writes: "Many young mathematicians are discouraged at their prospects for a successful career in mathematics because of decreased research funds, poor salaries and the shortage of openings in universities. Several I know are actively looking for jobs in other fields where they can expect much better treatment economically, and even the top departments are finding it increasingly difficult to attract qualified graduate students."

What is the research support picture in the mathematical sciences? One can glean an indication from Tables 1 and 2.

Lest the disparity in funding exhibited in Table 2 be totally attributed to differences in the numbers of professionals in the various areas, we note that in January 1980 the numbers of full-time

TABLE 1

	Percentage of Research in Universities Not Sponsored November 1978-October 1979
Engineering	16
Environmental Sciences	16
Life Sciences	13
Mathematical/Computer Sciences	59
Physical Sciences	22

Source: "Activities of Science and Engineering Faculty in Universities and 4-Year Colleges 1978-1979; Surveys of Science Resources Series NSF Final Report," NSF 81-323, page 2.

TABLE 2

	Estimated Federal Obligations for Research* Performed at Universities and Colleges, Fiscal Year 1981
Engineering	$ 350,208
Environmental Sciences	$ 332,063
Life Sciences	$2,088,893
Mathematics	$ 55,906
Physical Sciences	$ 511,638

*"Research" does not include "R&D Plant," which is defined as follows: "R&D plant (R&D facilities and fixed equipment, such as reactors, wind tunnels, and radio telescopes) includes acquisition of, construction of, major repairs to, or alterations in structures, works, equipment, facilities, or land, for use in R&D activities at federal or non-federal installations. Excluded from the R&D plant category are expendable equipment and office furniture and equipment. Obligations for foreign R&D plant are limited to federal funds used in support of foreign research and development." Ibid, page 2.

Source: "Federal Funds for Research and Development, Fiscal Years 1979, 1980, 1981," NSF 80-318, page 117.

scientists and engineers at doctorate-granting institutions in the various areas were as follows:

Engineering - 20,511
Life Sciences - 93,309
Physical Sciences - 16,845
Environmental Sciences - 5,891
Mathematical Sciences - 9,146*

In light of these data, it is not surprising that departmental chairmen speak of discouragement and of deterioration of the morale of young mathematicians. If it is in the national interest to maintain a healthy and vigorous mathematical sciences research enterprise, then it is imperative that conditions contributing to this deterioration be altered. We discuss the more important problem areas, and for each area estimate the dollar cost of alleviating the problem.

Postdoctoral Positions in the Mathematical Sciences

Table 3 (based upon data in "Academic Science: Graduate Enrollment and Support for 1980," NSF 81-330, Table A-30) gives dramatic evidence of the disparity in the numbers of postdoctorates in various sciences and in engineering.

The excellence of science in the United States today derives from postdoctoral opportunities in the past. Clearly, if the current postdoctoral pattern in the mathematical sciences persists, we will jeopardize the quality of the mathematical sciences at our leading universities in the years to come. To be sure, the pool of outstanding candidates for tenure positions at the leading three or four universities will be large enough. However, one can expect a serious drop in the quality of candidates at the next five ranking universities, and an even more serious drop at the next ten and twenty.

There is a clear need to provide a significant number of outstanding recent Ph.D.s in the mathematical sciences with the opportunity to devote full-time to research in association with a major scientific figure of their own choosing. The most creative future mathematicians in the United States will emerge from this group, and it must be nurtured. We are not advocating a large move away from teaching, the traditional mode of mathematical support. Typically, all graduate students will do some teaching while preparing for their doctorate. However, we are suggesting a small shift to allow promising young investigators a few years after their degrees to develop their research talents.

The Panel estimates that there should be an additional 120 postdoctoral appointments each year, each appointment being for two years.

---

*Data from "Academic Sciences: Scientists and Engineers," NSF 81-307, Table B-5. These figures include both research and non-research scientists and engineers.

TABLE 3

	Number of Postdoctorates in All Graduate Institutions: 1980
Engineering	981
Environmental Sciences	311
Life Sciences	11,715
Mathematics	143
Physical Sciences	4,261

At a cost of $25,000 per appointment per year, this amounts to an increment of $6 million per year. It is also essential, in the view of the Panel, that there be flexibility in the nature of the postdoctoral support. A variety of modes should be used in offering it: fellowships in NSF, DOE, DOD; institutional support at major centers and research institutes; enhancement of grants by providing support for postdoctoral positions.

Research Grants in the Mathematical Sciences

Table 4 indicates many categories where funding for the mathematical sciences is markedly insufficient. The amount of dollar support available for graduate students in the mathematical sciences is inadequate. We illustrate this with a very particular example. A member of this Panel recently directed two Ph.D. dissertation students, one in the mathematical sciences, the other in another science. The non-mathematics student was federally supported, and his sole task was to devote his time to study and research. The mathematics student, on the other hand, was university supported and, in addition to study and research, had to devote his time to grading papers, teaching, registering students, and holding office hours. The difference in treatment did not go unnoticed by these students nor, assuredly, did others fail to notice it. Mathematics students need the opportunity for a year or two of uninterrupted research during their graduate study to fully develop their research abilities, as students in most other fields of science do.

Along with decreased research support there is a rapid increase in the demand for mathematics courses in universities. Those who enter and remain in the mathematical sciences must necessarily teach more and devote less time to research. This is now characteristic of the mathematical sciences but not of most of the other sciences.

Because of this dearth of support and less than optimal conditions for the acquisition of knowledge in mathematics, many talented students select other areas for study. As a consequence, the quality and excellence of the graduate student body in the mathematical sciences are diminishing, and, as with the postdoctorates, the implications for the future excellence of the discipline in the United States are ominous.

TABLE 4 Bugetary Categories of NSF Awards: FY 1981

	Mathematical Sciences		MPS Directorate[a]	
Budgetary Constraints	Amount	Percentage of Total	Amount	Percentage of Total
Personnel:				
Senior Personnel	$11,710	41.6	$ 23,253	10.2
Postdoctoral Associates	1,699	6.0	17,068	7.5
Graduate Students	1,189	4.2	26,611	11.7
Other Personnel Costs[b]	2,538	9.0	29,027	12.7
SUBTOTAL: Wages, Salaries, and Fringe Benefits	17,136	60.8	95,959	42.1
Permanent Equipment	50	.1	38,550	16.9
Other Direct Costs[c]	2,022	7.2	41,141	18.0
Indirect Costs	8,972	31.9	52,370	23.0
TOTAL: All Budgetary Categories	$28,180	100.0	$228,020	100.0

[a] Excludes mathematical sciences section, MCS.

[b] Includes undergraduate students, secretarial-clerical, other professionals, technicians, and fringe benefits.

[c] Includes domestic and foreign travel, materials and supplies, publication costs, consultant costs, and computer costs.

Implicit in Tables 1 and 4 is that a substantial amount of support for research in the mathematical sciences must be contributed by universities in the form of reduced teaching loads, secretarial services, travel costs, and other aids. This support, as we have noted, is crumbling.

It is the view of the Panel that an additional 300 graduate students should be supported per year. At a cost of $20,000 per student (including indirect costs) this amounts to an increment of $6 million per year. In addition, the amount allocated in federal mathematical sciences research grants for secretaries, travel, and publication should be increased by $3,000 per individual investigator. This amounts to an increment of $5.4 million per year.

Percentage of Active Research People Supported

We saw in Table 1 that, in fiscal year 1979, 59 percent of the mathematical sciences research done in universities was unsponsored. Since then, the situation has deteriorated even further: *approximately 200 active researchers doing high-quality work have lost support during the past two years. Moreover, the research of many excellent new Ph.D.s in the mathematical sciences--more than 86 percent--goes unsupported*. We are not capitalizing on the investment made in the development of mathematical scientists.

To stem the decay and to put a measure of vitality into the mathematical sciences research enterprise require the allocation of sufficient funds to support an additional 500 researchers. At a cost of $20,000 per researcher, including indirect costs, this amounts to $10 million per year.

State of the Infrastructure

The stresses addressed thus far pertain primarily to those that affect the individual researcher or graduate student. Severe problems also exist in the area of communication and interaction between researchers. There is a paucity of mechanisms in the mathematical sciences for maintaining the vitality of researchers working at a distance from elite centers, for generating young people's interest in promising and important new subfields, and for informing and educating the research community about new ideas and results--particularly those ideas and results that lie at the boundary of two, or more, scientific disciplines. Those mechanisms that do exist are in disrepair. Support here is especially important because of the new opportunities made possible by the recent liaisons between mathematics and other sciences.

There has been substantial discussion in the mathematical sciences community concerning means of supporting the infrastructure of the research enterprise. In Table 5 we list some of these means and the amount of funds that the Panel feels will be required per year to support them.

Table 6 is a summary of the per annum dollar amounts discussed above (*not* listed in priority order).

TABLE 5

Summer schools, special years, mini-institutes	$5.3 million
Mid-level fellowships	3.6 million
Travel grants, senior research associate programs	4.5 million
Computer time and equipment	2.0 million
TOTAL	$15.4 million

TABLE 6

Postdoctoral positions	$ 6.0 million
Graduate students	$ 6.0 million
Operating expenses in grants	$ 5.4 million
Increase in number of grants	$10.0 million
Infrastructure (total)	$15.4 million
TOTAL	$42.8 million

SECTION 3 RECOMMENDATIONS

The mathematical sciences lie at the core of science, technology, and the national defense. An excellent mathematical sciences enterprise does not automatically produce excellence in science and technology and strength in the national defense, but one cannot have quality and strength in the latter without quality and excellence in the mathematical sciences.

The cost to the federal government of ensuring excellence in the mathematical sciences is relatively very small, the leverage of the dollars invested very large, the ratio of benefit to cost enormous.

In times of economic uncertainty or stress, the implementation even of programs with high benefit-to-cost ratios is sometimes delayed. When this occurs, one must also look at the "disbenefit" associated with delay. The current exciting opportunities will not be fully exploited in the United States; the erosion of excellence in the mathematical sciences will accelerate; the disbenefit to our nation of not doing something now is too large to allow the erosion to continue.

The Panel makes the following recommendations:

1. The federal dollar allocation for research in the mathematical sciences should be increased over the next three years by approximately 60 percent--i.e., there should be a total increment over the next three years of $42.8 million (in 1982 dollars).
2. The increments should be allocated in proportions deemed appropriate to the NSF, the DOD agencies, and the DOE.
3. The managers of mathematical science funding programs should have sufficient flexibility and freedom to choose the areas of research to be supported and the mechanisms for support.

The Panel recognizes that it has no formal status as representative of the mathematical community. No single group does. Nevertheless, because it realizes that increased support may be slow in coming, despite the emergency, the Panel tried to give priority to its recommendations as follows.

• If there is no increase, we recommend no changes. The mathematical community, after deliberation, has recently reallocated the resources available to it. It shifted monies into an alternative mode of research support and cut down on the number of individual research grants. It will take time to absorb these changes. We must emphasize again, however, that no increase in support will spell disaster for all but a few of the top mathematical research centers.

• If there is a 10 percent increase ($5.7 million), we recommend that the increment be disbursed approximately as follows:

Allocation	Use
$1.5 million	Postdoctoral Positions (in a variety of modes)
$1.2 million	Increasing the Operating Expense Allocations in Grants
$0.5 million for each of the following:	Mini-Institutes
	Senior Research Associate and Visiting Positions in Grants
	Mid-level Fellowships
	Graduate Student Support
	Equipment and Computer Time
$0.25 million for each of the following:	Increasing the Number of Grants
	Travel Grants, Special Years, etc.

• If there is a 20 percent increase ($11.5 million), we recommend that the increment be disbursed approximately as follows:

Allocation	Use
$2.0 million	Increasing the Operating Expense Allocations in Grants
$2.0 million	Postdoctoral Positions (in a variety of modes)
$1.75 million	Increasing the Number of Grants
$1.0 million	Senior Research Associate and Visiting Positions in Grants
$1.0 million	Graduate Student Support
$0.75 million for each of the following:	Mini-Institutes
	Mid-level Fellowships
	Research Institute
	Equipment and Computer Time
	Travel Grants, Special Years, etc.

- If there is a 50 percent increase ($28 million), we recommend that the increment be disbursed approximately as follows:

Allocation	Use
$4.8 million	Increasing the Number of Grants
$4.6 million	Postdoctoral Positions (in a variety of modes)
$4.3 million	Increasing the Operating Expense Allocations in Grants
$3.5 million	Senior Research Associate and Visiting Positions in Grants
$3.2 million	Graduate Student Support
$2.2 million	Mini-Institutes
$1.7 million	Mid-level Fellowships
$1.4 million	Equipment and Computer Time
$1.4 million	Research Institute
$0.9 million	Travel Grants, Special Years, etc.

- If there is an 80 percent increase ($42.8 million), we recommend that the increment be disbursed approximately as follows:

Allocation	Use
$6.0 million	Postdoctoral Positions
$6.0 million	Graduate Students
$5.4 million	Operating Expenses in Grants
$10.0 million	Increasing the Number of Grants
$5.3 million	Summer Schools, Special Years, Mini-Institutes
$3.6 million	Mid-level Fellowships
$4.5 million	Travel Grants, Senior Research Associate Programs
$2.0 million	Computer Time and Equipment

REPORT OF THE RESEARCH BRIEFING PANEL ON ATMOSPHERIC SCIENCES

Research Briefing Panel on Atmospheric Sciences

Charles L. Hosler, Jr. (Chairman), Pennsylvania State University
Richard A. Anthes, National Center for Atmospheric Research
Louis J. Battan, University of Arizona
Ralph J. Cicerone, National Center for Atmospheric Research
Peter V. Hobbs, University of Washington
John E. Kutzbach, University of Wisconsin, Madison
C. Gordon Little, NOAA Wave Propagation Laboratory
Michael B. McElroy, Harvard University
Norman A. Phillips, National Meteorological Center (NOAA)
Joseph Smagorinsky, Princeton University
Verner E. Suomi, University of Wisconsin, Madison

Staff

John S. Perry, Executive Secretary, Board on Atmospheric Sciences and Climate
Allan R. Hoffman, Executive Director, Committee on Science, Engineering, and Public Policy

REPORT OF THE RESEARCH BRIEFING PANEL ON ATMOSPHERIC SCIENCES

SUMMARY

Advancing scientific understanding and technology have put us on the threshold of dramatic improvements in our ability to predict the future state of the atmosphere. At this time, special opportunities for scientific advances, which also involve important social concerns, are presented by three areas of the atmospheric sciences: small-scale severe weather systems, atmospheric chemistry, and climate.

This brief report recommends specific actions that should expand our knowledge and predictive capability in each of these areas. The recommendations are based upon the scientific merits of the various actions and do not consider details of execution and cost. Advances in these three areas will improve human productivity and efficiency, protect life and property, and enhance the quality of life. (See Table 1.)

INTRODUCTION

The global atmosphere forms an indivisible, unified object of study: space-borne observation/communications platforms to view the earth as a whole and cooperative programs among all the world's nations are fundamental and indispensable prerequisites for meaningful research. Thus, while the atmospheric sciences, like other branches of the scientific enterprise, continually present challenging problems and opportunities, their progress depends to an unusual extent upon multi-disciplinary investigations, a global supporting infrastructure of space- and ground-based observing/monitoring programs (largely justified by operational requirements), and multinational activities of many kinds.

Man's influences on certain atmospheric processes are now comparable in some respects to nature's. By release of carbon dioxide, industrial effluents, and entirely new substances such as the chlorofluorocarbons, we may be changing the radiative and chemical nature of the atmosphere and thus affecting its ability to maintain a stable climate and a habitable world for humans.

TABLE 1 High-Leverage Investments in the Atmospheric Sciences

Problem Area	Recommended Action	Probable Benefits
1. Small-scale Severe Weather Systems	Use near-term technological enhancements (Doppler radar and wind profiling, data management and and display) as a basis for planning long-term mesoscale research	Immediate improvements in forecasts/warnings; improved aviation safety; basic understanding; accelerated progress in research; improved prediction techniques
2. Atmospheric Chemistry:		
a. Acid Rain	Conduct field studies on dry deposition, in-cloud processes, and remote-region sources	Cost-effective and environmentally sound control strategies
b. Tropospheric Chemistry	Support a comprehensive research program, including data base	Air quality protection and climate prediction
c. Stratospheric Ozone	Support the Upper-Atmosphere Research Satellite (and associated research)	Ozone layer protection
3. Climate Variability and Change		
a. Long-term Climate Change	Study the circulation of the World Ocean and its interaction with the atmosphere and climate	Predictive models for long-term climate change
	Measure and model trace gas concentrations	Assessment of effects on radiation balance and climate
b. Short-term Climate Variations	Support El Nino/Southern Oscillation studies	Seasonal/interannual predictions
	Conduct prediction studies with general circulation models	Monthly/seasonal predictions

Through daily weather, through the slow march of climate, through effects and control of pollution, the atmosphere influences myriad daily activities and decisions throughout our complex society, all of which contribute to our nation's efficiency and productivity. The atmosphere's constituency thus includes not only pilots, farmers, commodity dealers, and builders, but every citizen. To ensure our continued welfare in a variable and vulnerable atmosphere, we must acquire better fundamental understanding of the atmosphere and an enhanced capability to predict its changes arising from both natural processes and human interventions.

SMALL-SCALE SEVERE WEATHER SYSTEMS

During the past decade there has been a growing consensus within the meteorological community that significant improvements can be made in the short-range forecasting (0-12 hours) of mesoscale--or stormscale--weather systems. In terms of size and duration, these lie between individual clouds and the larger storms depicted on the conventional weather map. Although they tend to elude the conventional observing/forecasting system, they produce much of our weather and most of our severe weather: tornadoes, flash floods, hailstorms, snowstorms, and strong winds. Optimism about our ability to forecast these systems is based on recent advances in both technology and understanding. Technological improvements include satellite systems (which provide information on cloud cover and temperature and humidity soundings), various types of radars (which can provide continuous and quantitative information on winds, precipitation, and--through pattern recognition--various weather phenomena), and electronic means for the rapid assimilation of this large quantity of information and its communication to the public in the form of timely forecasts. Scientific advances have been made in the understanding of such phenomena as frontal circulations and precipitation, severe mountain winds, and severe thunderstorms which produce hail, wind shear, and sometimes tornadoes.

Currently, much of this technology is not available to the public or private weather forecaster, although it has been well tested in the research community. Transfer of the technology to the forecasting sector should enable weather services, public and private, to make more timely and accurate predictions of stormscale weather phenomena. Such improvements would aid in protecting life and property, serve the national economy, and help meet defense requirements. Widespread employment of this technology would also provide a badly needed research data base and accelerate the progress of research in this area.

Long-Term Mesoscale Research

Two areas should be emphasized in long-term mesoscale research. First, we must improve our understanding of the physical processes that determine the growth, maintenance, and decay of mesoscale weather systems. This improved understanding can be achieved through appropriate special observational programs and theoretical research. Such knowledge is essential for the development and improvement of the conceptual and numerical models necessary to improve the forecasts and warnings of significant mesoscale weather.

The second major research objective is to develop methods of analyzing the massive data sets that would be available from the new measuring facilities. The development of analysis schemes and the communication systems required to process, transmit, and display the data are necessary both for short-range forecasts and warnings and for numerical weather prediction models. The recommendations outlined below will lead the way to more effective research in these areas, and will also begin to realize the improvements in short-range severe weather forecasting made possible by new technology.

Near-Term Technological Enhancements

Rapid exploitation of several technological advances could quickly accelerate our research progress and improve our forecasting capabilities. Doppler radar operating in the near-horizontal can detect the special motion patterns within thunderstorms that precede the formation of tornadoes. Also, the formation of weather-producing stormscale circulations, such as squall lines or heavy snow bands, can be detected from small changes in wind patterns before they have formed. Tests have demonstrated that Doppler radar observations can lead to earlier issuance of tornado warnings and an appreciable reduction in false alarms, thus enhancing public confidence in and response to warnings. Such radars can also warn of thunderstorm-generated severe downdrafts and wind shears that pose hazards to aircraft in landing and takeoff. Federal recognition of this potential already exists in the interagency NEXRAD program, involving the Department of Defense, the Federal Aviation Administration, and the National Weather Service. Funding for this program should be accelerated in order to keep all essential elements of the program moving on schedule. Full financial support of this system presents the most effective contribution to improved forecasting of severe weather that the federal government can now bring about. The routine use of Doppler radar will also increase scientific understanding of these destructive and erratic wind storms and the small-scale circulations hazardous to aviation.

- In the near term, the technological advances offered by Doppler radar should be exploited by accelerating funding of the interagency NEXRAD program.

Radar wind profiling systems using longer wavelengths can measure upper-air winds to considerable heights from naturally occurring refractive index variations. Radiometric temperature sounding systems can similarly obtain temperature and pressure data. The systems operate unattended and produce accurate and continuous observations in all weather conditions. The latter characteristics are completely new and highly valuable from both scientific and operational viewpoints. Continuous observations permit accurate tracking of significant characteristics of the flow of weather, such as fronts and wind shear lines.

The continuous availability of enroute winds for aircraft would also make possible significant fuel savings. In addition, dual-polarization Doppler radars can provide valuable information on cloud and precipitation processes for research purposes.

- An array of Doppler radar upper wind sounding systems should be deployed promptly in parts of the country most subject to violent weather to permit evaluation of their impact on weather forecasting and to aid research on mesoscale weather systems.

Both research and applications on the storm scale are currently limited by the coarse horizontal and temporal resolution of our present upper air data network and data handling systems. Radar can provide additional information, and sounding instruments on geostationary satellites are beginning to provide immense improvements in horizontal and temporal resolution. In addition, frequent images in the visible, the infrared and water vapor bands are becoming available from satellites. Interpreting this enormous flood of data has been likened to trying to drink from a fire hydrant without drowning. Over the last several years, however, demonstrations have shown that it is possible to digest this data flood in real time and to convert it into highly useful information for forecasting and research. To do this, the user must be provided with computer-based video displays and graphics capabilities. When this is done, the improvement in forecasting performance has been dramatic. Unfortunately, our existing system for communicating and displaying this information to the user falls far short of what is possible. The communications industry continues to develop an impressive number of new communications systems, including direct broadcast from satellites, that could serve to collect and disseminate weather information and support services to a wide range of users including researchers, forecasters, and the general public. This is an unusual opportunity to provide large improvements at reasonable cost.

These capabilities can be and should be made available in the near future for use by the weather services and the research community. Their widespread availability would provide immediate economic advantages and reductions in losses of lives and property. Moreover, they would provide the basic framework needed for a longer-term program of research on small-scale weather systems. In addition, the availability of state-of-the-art data management and display capabilities to the research community would significantly accelerate progress in other important areas such as climate.

- Federal agencies should promptly initiate actions to make advanced communications, data management, and display capabilities available both to the research community and to operational forecasters.

The above near-term technological enhancements would provide the basis for development of a long-term research program to improve our scientific understanding and operational capabilities in small-scale severe weather systems.

ATMOSPHERIC CHEMISTRY

The past decade has seen a revolution in the field of atmospheric chemistry. In the 1970s, chemists were primarily engaged in identifying problems and reacting to crises; in the 1980s they are seeking a new level of maturity for this area of science. They have learned that crisis assessment, in the absence of a deep understanding of the chemistry and physics of the natural atmosphere, is a hazardous exercise.

In a larger context, the earth is a planet characterized by change. Moreover, the human race has now achieved the ability to alter its environment on a global scale, with implications for food production, the quality of air and water, and the integrity of the global chemical cycles essential to life. In the past, human needs could be met by the expansion of frontiers, by land clearance, by the application of chemical fertilizers and pesticides, by irrigation, and through the exploitation of energy resources harvested from the sun by the biosphere millions of years ago. An understanding of the overall system is essential if the human race is to live successfully with global change.

Several current issues relating to atmospheric chemistry are singled out below for special near-term emphasis.

Acid Rain

Many of the general problems facing atmospheric chemistry are illustrated by the acid rain problem. Unlike Europe, no routine monitoring of the chemical properties and acidity of precipitation has been conducted in North America until very recently, and the effects of increased industrialization on the acidity of precipitation have been largely undocumented. Similarly, little is known about the chemical composition and acidity of "natural" (i.e., "clean" background) rainfall in remote regions of the world, which sometimes exhibit unexpected acidity. Since the cost of controlling industrial emissions is large, and the economic consequences could be great, the effects of control strategies on acid rain should be carefully assessed scientifically.

The principal contributors to acid rain are compounds of sulfate and nitrate, both of which are released by human activities. These compounds may appear in precipitation through two mechanisms: "scavenging" of sulfate and nitrate compounds already present in the atmosphere and/or "*in situ*" production through incorporation of trace gases into cloud and precipitation droplets followed by chemical reactions within the droplets to form sulfates and nitrates. Acidic material also reaches the ground through dry deposition. Establishment of the dominant mechanism ("scavenging," "*in situ*" production, or dry deposition) will dictate, in part, appropriate control strategies; it might even indicate that currently available control techniques would have relatively little effect on the occurrence of acid rain.

- *Three essential investigations should be undertaken in an attack on the problem of acid rain: Field measurements of dry deposition for acidic gases, e.g., SO_2 and HNO_3; field studies of in-cloud chemistry* and cloud physics from aircraft or at mountain laboratories; and assessment of natural sources of acidity in remote regions.

Tropospheric Chemistry

An understanding of the acid rain problem will require a broad assessment of tropospheric chemistry, not simply on a local but also on a global scale. For example, a major gap has developed in the measurement and understanding of tropospheric ozone. Available data and theory indicate that commercial aircraft operations can increase tropospheric ozone. Photochemical pathways to ozone production in the troposphere should be simpler than those in urban smog, and measurements of the relevant variables should be easier. Coordinated measurements are needed, techniques need to be developed, and global measurements from space will eventually be required to provide a data base for global tropospheric models.

More generally, tropospheric chemistry has a major impact on the quality of life for the planet as a whole, e.g., air quality and climate stability. The natural state of the tropospheric system, including its interaction with the global scale biosphere, must be defined to provide a basis for assessment of such issues as acid rain. A long-term integrated program of observations and research is now being defined by a study group of the National Research Council.

- Acquisition of a data base and a level of understanding sufficient to define the natural state of the troposphere should be a major goal for an integrated program of tropospheric research. To this end a program of research should be mounted to develop data on the distribution of important gases such as ozone, and a strategy should be implemented to improve our understanding of chemical cycles, with definition of important biospheric sources and sinks for key species.

Stratospheric Ozone

Scientific interest is now aimed at improving basic understanding of the stratospheric photochemical system. It is possible that the redistribution of stratospheric ozone caused by man's activities will be as important environmentally as the change in the total amount of ozone. Fuller quantitative understanding of man's impact on the ozone layer is needed to provide a sound basis for future decisions by industry and government.

Observations are needed to test available models of stratospheric chemical and dynamic processes and to guide the development of future models. Strategies for an orderly study of the stratosphere have been developed involving all segments of the concerned scientific communities--atmospheric chemists, laboratory chemists, meteorologists, and modelers. The research is interdisciplinary and international in scale, and it demands the collaboration of government, academia, and the industrial sector.

- The Upper Atmospheric Research Satellite (UARS) is a central element of this nation's stratospheric research program, and a commitment to this satellite program should be made at the earliest possible time.

CLIMATE

In the past dozen years, the attention of atmospheric and oceanic scientists has increasingly turned to problems of climate. Quantitative study of past climates has revealed the long-term variability of climate, while the Sahel drought and the grain trade dislocations of the early 1970s showed our vulnerability to changing climate. Numerical models and field observations (e.g., the measurement of increasing atmospheric concentrations of carbon dioxide and some other radiatively active trace gases) raised concern that our own actions might be producing damaging and irreversible changes in climate. However, it was realized that both prediction of natural climate variations and assessment of man-made changes require as an indispensable prerequisite a basic understanding of the workings of the global climate system comprising the oceans, the land surface, the ice and snow masses, and the atmosphere, and the ability to construct quantitative models based upon this understanding. These concerns prompted a worldwide surge of research on the global climate system and the organization of coordinated national and international programs. They also focused interdisciplinary research effort on critical problems such as the increase in atmospheric carbon dioxide. These endeavors are well under way and should proceed unabated. However, they have already revealed several opportunities that merit exploitation.

World Ocean Circulation

The World Ocean, covering 70% of the planet's surface, is an integral component of the climate system and provides its long-term memory. The dynamics of the oceans, however, are only poorly understood in comparison to those of the atmosphere, impeding our ability to model the coupled combined systems. Hence, we are unable to determine adequately the important thermal and chemical interactions between the ocean and the atmosphere. As a result, we cannot predict climate beyond a season or elucidate completely the sensitivity of climate to a CO_2 increase.

A fundamental necessity is to establish observationally the three-dimensional structure of the circulation of the World Ocean, its seasonal variation, its interannual variability, and its interactions with the atmosphere. Only then can we establish definitively the relative role of the ocean in storing and transporting heat in the climate system, stimulate modeling improvements, and supply a validation base for climate simulation studies. Such advances would be of equal benefit to the discipline of oceanography.

New technology for ocean measurements makes it possible to conceive of an observational experiment to determine the circulation of the World

Ocean. Satellite-borne radar altimeters can determine the sea-surface topography, while scatterometers coupled with other observations provide measures of wind stress. Acoustic tomography promises ocean-scale density measurements, and measurements of man-made tracers directly reveal long-term circulations. These and other observations will supplement the ongoing global observing system.

- Investments should be begun now to ensure the availability of long lead-time satellite-borne observing and communications capabilities (typified by the TOPEX program) to permit initiation of a world ocean circulation experiment by the end of this decade.

Trace Gases

Much attention has been focused on projected global warming due to increased atmospheric CO_2. However, the corresponding effects of other radiatively active trace gases are additive to those of CO_2 and can be significant. Concentrations of nitrous oxide, the chlorofluorocarbons, methyl chloroform, and methane have increased worldwide, and there is circumstantial evidence that tropospheric ozone is also increasing. These and other gases absorb infrared energy in otherwise transparent regions of the infrared spectrum and thus block heat from leaving the earth.

Monitoring the concentrations of these gases (except for ozone) is economical and feasible. Vertical profiles with good latitudinal coverage are needed for nitrogen oxides, ozone, some hydrocarbons, water vapor, and carbon monoxide, among others, as a basis for theoretical and modeling improvements.

- Field and laboratory measurements should be taken to provide a firmer basis for our understanding and prediction of the behavior of the trace gases that influence climate.

El Nino, the Southern Oscillation, and Interannual Predictability

A large and growing body of evidence suggests that the tropical oceans influence the global circulation of the atmosphere on time scales from a season to about two years, and provides encouraging possibilities for prediction of seasonal and interannual climate fluctuations.

The Southern Oscillation encompasses a large variety of phenomena throughout the global atmosphere and ocean that are known to be correlated with the large-scale oscillations of south Pacific surface pressure to which the term was first applied. These phenomena include large increases in water temperature off the Peruvian coast (El Nino), equatorial sea-surface temperature anomalies, variations in the Indian Monsoon, and seasonal climate irregularities over North America. A full cycle of the Southern Oscillation endures over about a two-year period, and is initiated aperiodically at intervals of two to ten years. It is the largest coherent signal in short-term climate variability.

Much of what has been learned recently about the interannual variability of the tropical ocean and the global atmosphere has come from existing observational sources and climate models, and the implications of these findings for climate predictability should be exploited promptly. Already, simple empirical methods based on these phenomena are being applied to produce crude seasonal forecasts over North America.

Further progress will require a better understanding of atmospheric and, especially, oceanic mechanisms. A more complete description of the chain of events is needed. What are the energetics of the cycle? What triggers it? Is there a beginning to the sequence? What determines the length of the cycle? Why is the frequency of occurrence so irregular?

A strong base for productive research exists in a group of Pacific-oriented oceanographic programs completed, in progress, or already planned. Study of this problem is now under way, and a national plan for a comprehensive research program in both the atmosphere and the ocean will be completed before the end of 1982. International interest and participation is already evident.

- A concentrated program of observation and research should be mounted in the early to mid-1980s in strong and continuing support of El Nino/Southern Oscillation studies.

Monthly and Seasonal Prediction

Recent research with global atmospheric general circulation models designed for prediction in the 5- to 15-day range indicate that in some instances atmospheric events may be predictable in the 30-day range. In particular, blocking episodes* and their attendant dramatic regional temperature and rainfall anomalies and shifts in storm tracks have been successfully predicted. There is some evidence that this enhanced predictability could be extended to the seasonal range. The investigation of atmospheric predictability was a prime objective of the Global Atmospheric Research Program, in particular of the 1979 Global Weather Experiment. This highly successful international endeavor yielded a unique body of detailed data on the time-dependent behavior of the atmosphere. Success in this research will require the sustained availability of substantial research funds.

- Comprehensive observational and theoretical studies should be conducted to discover the physical reasons for persistent circulation states and to determine whether there are any other circulation characteristics that may also be predictable in the monthly/seasonal time range.

---

*The term "blocking" refers to the development of strong and almost stationary circulation patterns that persist over large regions for extended periods, i.e., a few days to several weeks.

REPORT OF THE RESEARCH BRIEFING PANEL ON ASTRONOMY AND ASTROPHYSICS

Research Briefing Panel on Astronomy and Astrophysics

George Field (Chairman), Harvard-Smithsonian Center for Astrophysics
Geoffrey Burbidge, Kitt Peak National Observatory
George Clark, Massachusetts Institute of Technology
Arthur Code, University of Wisconsin
Carl Fichtel, NASA Goddard Space Flight Center
William Hoffman, University of Arizona
Kenneth Kellermann, National Radio Astronomy Observatory
Jeremiah Ostriker, Princeton University
Tobias Owen, State University of New York, Stony Brook
Harlan Smith, University of Texas, Austin
Harvey Tanabaum, Harvard-Smithsonian Center for Astrophysics
Arthur Walker, Stanford University

STAFF

Paul Blanchard, Rapporteur (Harvard University)
Bruce Gregory, Executive Secretary, Space Science Board
Allan R. Hoffman, Executive Director, Committee on Science, Engineering, and Public Policy

REPORT OF THE RESEARCH BRIEFING PANEL ON ASTRONOMY AND ASTROPHYSICS

I. ASTRONOMY AND ASTROPHYSICS FOR THE 1980's

We live in an extraordinary age for astronomy, unmatched since the time of Galileo. The discoveries of the past 20 years, made from both ground-based and space observatories have radically changed our understanding of the Universe. Quasars, x-ray sources, pulsars, polyatomic molecules in interstellar space, and the cosmic microwave background radiation were discovered in the 1960's. The 1970's added discoveries of accretion of matter by neutron stars, observational evidence for a stellar black hole, intergalactic gas at 100 million degrees, apparent expansion at greater than the speed of light in quasars, volcanic activity on Io, the origin of high-velocity solar-wind streams in "holes" in the solar corona, rings around Uranus and Jupiter, observational evidence for gravitational radiation from a binary star, gamma-ray bursts, and the gravitational refraction of light from quasars by intervening galaxies.

Through federal support, U.S. astronomers have remained at the forefront of astronomy and astrophysics during this period, maintaining a position of world leadership established early in this century. Spectacular advances in x-ray, gamma-ray, infrared, and ultraviolet astronomy--supplementing the more established techniques of optical and radio astronomy--have opened virtually the entire electromagnetic spectrum to cosmic observations. The U.S. scientific community stands ready to exploit this opportunity with present and planned facilities of unparalleled power and scientific promise.

Special "decade reviews" carried out by the National Academy of Sciences (NAS) have periodically surveyed the status of the field and have advanced recommendations for future research programs. The first two such surveys, completed in 1964 and 1972, had a powerful impact on the conduct of U.S. research during the period 1960-1980. The recommendations of the third such decade review are contained in the recent report of the Astronomy Survey Committee, Astronomy and Astrophysics for the 1980's (National Academy Press), Washington, D.C., 1982; see Appendix A). Timely response to these recommendations will permit U.S. scientists to maintain leadership in the most promising areas of astronomy and astrophysics during the 1980's.

The Astronomy Survey Committee report took three years to complete and engaged over 130 scientists. Its final recommendations were highly selective: of many dozens of projects considered, only a small number were recommended for new funding for the 1980's (see Appendix B for a summary taken from the 1982 report). The National Academy of Sciences has accepted these recommendations as the consensus of the U.S. astronomical community, and they have also received the official endorsement of the American Astronomical Society.

The present Briefing Panel has based its recommendations for FY 1984 entirely on the recently completed Survey Committee report, and they follow the same format as does the report. The Survey Committee first stressed the importance of Approved, Continuing, and Previously Recommended Programs for progress in research during the 1980's. The Committee next emphasized that Prerequisites for New Research Initiatives are essential to the success of present and planned major research facilities. Finally, the Committee recommended the funding of New Programs for the 1980's in three categories: Major, Moderate, and Small. Throughout its study, the Survey Committee strove to be fiscally realistic: the overall program recommended for the 1980's is roughly comparable in scale with that actually carried out during the 1970's on the basis of the recommendations of the previous decade review.

From the projects recommended for the 1980's in the Survey Committee report, the Briefing Panel has selected for recommendation here only those that are particularly timely for FY 1984 budget action. (It is important to recall again that the programs recommended by the Survey Committee themselves reflected a careful selection from a great many proposals made to that Committee.) Favorable action by the government in FY 1984 on the programs recommended here will constitute an important step toward ensuring the vitality of U.S. astronomical research in the decade ahead. The Panel calls particular attention to the need for FY 1984 commitments to begin two important projects: the Shuttle Infrared Telescope Facility (SIRTF) within the NASA Spacelab program and a Very Long Baseline Array (VLBA) of radio telescopes within the program of the NSF Astronomy Division.

II. THE RECOMMENDATIONS

The recommendations of the Briefing Panel for FY 1984 are summarized in this section in the same format as in the Survey Committee report. They are discussed in more detail in Section III.

Approved, Continuing, and Previously Recommended Programs

Shuttle Infrared Telescope Facility (SIRTF). This major Spacelab facility is the only space-astronomy project in this category that has not yet been started. SIRTF will achieve thousandfold gains over present telescopes in the important thermal-infrared region of the spectrum. The Panel recommends that NASA re-establish its original strong commitment to a vigorous Spacelab program by immediately implementing SIRTF through a Phase B study, which will require an augmentation to the Spacelab program of $2 million in FY 1984 and $2.5 million in FY 1985.

Prerequisites for New Research Initiatives

Instrumentation and Detectors, Theory and Data Analysis, and Computational Facilities. The Survey Committee identified five Prerequisites which are essential to the success of major research programs but which are inexpensive by comparison. The Panel has selected the three Prerequisites stated above because funding them will immediately increase the efficiency with which observations are made and analyzed and the degree to which the resulting data can be understood on the basis of physical laws. The Panel recommends that an augmentation of $10 million/year above present funding levels, to be shared by NSF and NASA, be made available for these Prerequisites beginning in FY 1984.

NEW PROGRAMS

Major New Programs

Advanced X-Ray Astrophysics Facility (AXAF)

This facility, which was accorded the highest priority among all major new programs by the Survey Committee, will be the nation's first permanent x-ray observatory in space. The Panel recommends augmentation of Phase B design studies by $2.5 million in FY 1984 and by $5 million in FY 1985 in order to ensure a new start for AXAF in FY 1986.

Very Long Baseline Array (VLBA) of Radio Telescopes

This ground-based facility, which was accorded the highest priority among major new ground-based programs by the Survey Committee (and second highest priority among all major new programs), will permit radio astronomers to map detail in cosmic radio sources with an angular resolution of 0.2 milliarcsecond. The Panel recommends that, in FY 1984, NSF commit to the construction of the VLBA and ensure its timely completion by making available $2-3 million for a final design study in FY 1984. Construction will require additional funding of $60 million spread over FY 1985-1987.

New Technology Telescope (NTT) of the 15-Meter Class

A ground-based NTT, which was accorded third highest priority among all major new programs by the Survey Committee, will make unique contributions to optical and infrared astronomy, particularly through optical and infrared spectroscopy of faint stars and galaxies and through high angular resolution studies in the infrared. The Panel recommends that design studies for the NTT, which were considered to be of the highest priority by the Survey Committee, be augmented to the level of $2 million in FY 1984 and in FY 1985 so that a final design for NTT can be selected by FY 1986.

Moderate New Program

Augmentation to the NASA Explorer Program

The Explorer program remains a flexible and highly cost-effective means to pursue new opportunities in space astronomy in nearly every part of the electromagnetic spectrum. The Survey Committee recommended an immediate and substantial augmentation to the Explorer budget in order to restore it to the healthy level of effort of 1970, which prevailed before the ravages of inflation. The Panel recommends that the Explorer budget be augmented by $20 million/year beginning in FY 1984 in order to accelerate the development and launch of three Explorer missions now in various stages of preparation: the Cosmic Background Explorer (COBE), the Extreme Ultraviolet Explorer (EUVE), and the X-Ray Timing Explorer (XTE).

III. JUSTIFICATION FOR THE RECOMMENDATIONS

Approved, Continuing, and Previously Recommended Programs

Shuttle Infrared Telescope Facility (SIRTF)

The programs recommended in the Astronomy Survey Committee's report were selected from research activities that were, at the beginning of the survey, candidates for implementation in FY 1983 and beyond. However, the Committee also emphasized "the importance of approved, continuing, and previously recommended programs to the progress of astronomical research during the remainder of the decade. The present Committee's recommendations take explicit account of such programs and build upon them" (Chapter 2, page 13; see also Appendix B of this Briefing Panel report). The Panel notes that SIRTF is the only space-astronomy project in this category that has not yet been implemented. The Survey Committee went on to say (Chapter 4, page 112--see also Appendix C to this Briefing Panel report), "The proposed SIRTF will be the cornerstone of research in infrared astronomy during the 1980's. The Astronomy Survey Committee joins with the Space Science Board's Committee on Space Astronomy and Astrophysics . . . in recommending this facility as the first major infrared telescope in space." The Panel enthusiastically supports this recommendation.

In support of its recommendation for SIRTF, the Survey Committee wrote: "SIRTF will permit investigations over the enormous range of wavelengths from 2 to 300 μm. For some important types of observations it will, because of its cryogenically cooled optics, yield a sensitivity gain of 1000 over the largest existing ground-based and airborne infrared telescopes; this gain in sensitivity is so large that it is not unreasonable to expect that SIRTF will make important and unexpected discoveries."

SIRTF will make it possible to observe the highly red-shifted light originating far in the past from very distant galaxies, so that we may study their early evolution. It will be able to detect very low-mass stars in nearby galaxies, which may account for the hidden mass in those galaxies, and to penetrate the obscuring dust of the dense cores of molecular clouds to study the process of star formation. SIRTF will also be able to investigate the thermal energy balance of the more distant planets and to greatly improve our knowledge of their atmospheric compositions.

Because of its wavelength range and sensitivity, SIRTF will in addition permit detailed spectroscopic studies of the infrared sources which will be discovered by the IRAS infrared survey satellite due to be launched in the very near future. Just as the 200-inch Hale telescope is used to investigate the stars and galaxies discovered by the 48-inch Schmidt telescope on Mt. Palomar, SIRTF will investigate new infrared sources discovered by IRAS.

A reaffirmation of NASA's original strong commitment to Space Shuttle research beginning in FY 1984 will be necessary to permit the timely development of SIRTF within the Spacelab program. With the flight of IRAS the development of cryogenically cooled optical systems for infrared space observations will have been completed and tested in space, thus removing a major barrier to the rapid development of SIRTF, a much more powerful and sensitive facility.

Prerequisites for New Research Initiatives

The Astronomy Survey Committee called attention to five Prerequisites for New Research Initiatives and recommended augmentations in each: "These Prerequisites are essential for the success of major research programs but are inexpensive by comparison." The Briefing Panel recommends that three of the Prerequisites receive augmentations in FY 1984.

Instrumentation and Detectors

Great strides in the development of new instruments and detectors were made during the 1970's, permitting enormous increases in observing efficiency, but so far only a handful of the telescopes at major observatories have been equipped with state-of-the-art instruments and detectors. Moreover, new generations of instruments and detectors must be developed in concert with the design and operation of future ground-based and space observatories. The application of state-of-the-art technology will greatly amplify the quantity and quality of the data obtained, at a small fraction of the cost of the major facilities that will use these instruments and detectors. Augmented funding for instrumentation will also permit accelerated development of a 10-meter submillimeter-wave radio antenna--the Astronomy Survey Committee's highest-priority recommendation for small new programs--by the NSF.

Theory and Data Analysis

The aim of astronomy is an understanding of the Universe, not simply the gathering of observational data. Vigorous programs of data analysis and complementary theoretical investigations are essential to achieve this aim, but both have been underfunded in the past. In particular, NASA has not adequately supported the theoretical analysis of data from space astronomy missions. Moreover, theoretical understanding of recent discoveries is important for optimizing the designs of future instruments. Increased support by NASA and NSF for theory and data analysis will greatly enhance the scientific return from both ground-based and space-astronomy facilities.

Computational Facilities

Modern instrumentation and detectors are highly dependent upon computers for control, data readout, and on-line processing. Because of the complexity of astronomical systems such as exploding stars and galaxies, theoretical investigations and data analysis cannot be carried out by purely analytic methods, and are therefore also dependent upon numerical calculations. The steady and rapid decrease in the cost of computing power opens up vast new opportunities for scientific advance. An augmentation to NSF and NASA funding of computational facilities will be a major step in achieving maximum effectiveness of new instruments and data-analysis programs.

NEW PROGRAMS

Major New Programs

Advanced X-Ray Astrophysics Facility (AXAF)

The United States pioneered world leadership in x-ray astronomy, beginning with the 1962 discovery of extrasolar x-ray sources by means of a brief rocket flight, through the first all-sky survey of x-ray sources carried out by the UHURU satellite in 1970, to the flight of High Energy Astronomical Observatory (HEAO) satellites during the period 1977-1981. As a result, x-ray observations have now attained an importance in contemporary astronomy comparable to those in other wavelength regions. In FY 1986, following completion of Phase B studies, NASA should begin construction of the Advanced X-Ray Astrophysics Facility (AXAF), the nation's first permanent x-ray observatory in space. The Astronomy Survey Committee accorded this project highest priority of all major new programs for the 1980's.

With up to 100 times greater sensitivity than any previous x-ray mission, AXAF will provide U.S. astronomers with a powerful and long-lived new capability. Working in concert with the Very Large Array and the VLBA in the radio region, and with Space Telescope in the visible and ultraviolet spectral regions, AXAF will be an important part of an ensemble of major U.S. facilities covering much of the electromagnetic spectrum and capable of observing objects simultaneously at different wavelengths; this capability will greatly increase the scientific value of the data obtained.

AXAF will be used to study faint x-ray sources in the farthest reaches of our Galaxy as well as individual high-luminosity sources in hundreds of other galaxies. High-resolution imaging, together with spectroscopic and polarimetric observations, will reveal the composition and dynamics of supernova remnants, galactic halos, and clusters of galaxies. Observations of remote galaxies and quasars will probe the effects of evolution in the early Universe.

Augmentations of funding for final AXAF design studies in each of FY 1984 and FY 1985 are necessary for the timely completion of Phase B studies.

Very Long Baseline Array (VLBA) of Radio Telescopes

The United States has led the world in the development of Very Long Baseline Interferometry (VLBI), a technique that permits the combination in a central computer of radio-frequency observations from widely separated observing sites to produce images of cosmic objects with extraordinary angular resolution. The Very Long Baseline Array (VLBA) will include about 10 antennas, placed in Alaska and Hawaii as well as across the continental United States. This facility will map cosmic radio sources of very small angular size with a fineness of detail equal to that which could be achieved by a single radio dish as large as an entire continent. The angular resolution achieved--0.2 milliarcsecond--is equivalent to being able to read, from Los Angeles, a Post headline in Washington, D.C. Among many applications of profound importance, this instrument will probe the small-scale structure surrounding the enigmatic energy sources in the cores of quasars and active galactic nuclei and will directly determine the distance scale within our Galaxy with unprecedented accuracy (see Appendix D).

The National Science Foundation is currently completing a peer review of the VLBA, on the basis of which the Panel feels that NSF should be in a position to make a positive commitment to this important program. Accordingly, the Panel recommends that in FY 1984 NSF commit to construction of the VLBA and ensure its timely completion through award of $2-3 million for a final design study in FY 1984. The Array utilizes proven technology, and its costs may be accurately estimated. Construction of the VLBA will require additional funding of $60 million spread over FY 1985-1987. However, it will be critically important for NSF to construct and operate the VLBA without a further erosion of programmatic base funds within the NSF Astronomy Division.

Approval of the VLBA will mark the end of a long hiatus in the construction of new ground-based astronomical facilities by NSF. The 1970's saw the completion of the Very Large Array radio telescope near Socorro, N.M., but construction of a 25-meter millimeter-wave radio telescope--also recommended by the 1972 decade review committee--was deferred for a number of years for fiscal reasons and is not at present being considered for funding in the form originally proposed. During this period, other nations have taken over the lead from the United States in the important field of millimeter-wave studies of atoms and molecules in space. The field of VLBI, which was also pioneered in the United States, represents an outstanding opportunity for maintaining leadership in this exciting field of science.

New Technology Telescope (NTT) of the 15-Meter Class

In considering this project, the Astronomy Survey Committee wrote: "The Committee finds the scientific merit of this instrument to be as high as that of any other facility considered and emphasizes that its priority ranking (No. 3 among major new programs) does not reflect its scientific importance but rather its state of technological readiness. The design studies needed before NTT can be constructed are of the highest priority and should be undertaken immediately." Thus, the Committee emphasized the importance of an immediate start toward the ultimate goal of NTT. The NTT has been studied at a low level for the past several years; because of the technological challenge of the NTT project, an augmentation of these studies to the level of $2 million/year in each of FY 1984 and 1985 will be required to permit a final design to be selected by FY 1986.

A ground-based telescope in the 15-meter class will make powerful contributions to optical and infrared astronomy, particularly through spectroscopic observations of faint stars and galaxies that cannot practicably be carried out with present instruments and through infrared observations with high angular resolution. New technology developed during the 1970's has made it feasible to build telescopes on this scale at a cost far lower than would have been possible only a few years ago. The capabilities of NTT will complement those of Space Telescope, which has higher angular resolution but less light-gathering power in the optical region of the spectrum, and of SIRTF, which has higher sensitivity but lower angular resolution in the infrared spectral range. The spectroscopic studies required to follow up on the discoveries by ST can be carried out only with an instrument having the aperture of NTT.

Moderate New Program

Augmentation to the NASA Explorer Program

Among moderate new programs recommended for the 1980's, the Astronomy Survey Committee listed an augmentation to the NASA Explorer program as the highest priority: ". . . the flight of new Explorer missions has in recent years fallen much below the rate needed for healthy advance. The rate will decline even more drastically during the early 1980's if present budget levels are not increased. The Astronomy Survey Committee thus recommends an immediate and substantial augmentation to the Explorer program to restore it to at least the real level of effort of 1970." An augmentation of $20 million in FY 1984 represents an important first step in response to this recommendation. Such funding will permit the accelerated development and launch of three Explorer missions currently in various stages of preparation.

1. The Cosmic Background Explorer (COBE), now under development, will test the big-bang model of the Universe by extremely precise measurements of the cosmic microwave background radiation.
2. The Extreme Ultraviolet Explorer (EUVE), at present under study, will conduct the first all-sky survey of EUV sources, closing an important gap in our knowledge of the cosmic electromagnetic spectrum.
3. The X-Ray Timing Explorer (XTE), also now under study, will permit studies of the variability of x-ray sources on time scales from milliseconds to years, providing new constraints on models of neutron stars, accretion disks around collapsed stars, x-ray bursters, and quasars.

Appendix A

[Reprinted from Science, August 20, 1982, Volume 217, No. 4561, p. 691. Courtesy, American Association for the Advancement of Science.]

Astronomy and Astrophysics in the United States

A recently issued report[*] on astronomy begins: "Nature offers no greater splendor than the starry sky on a clear, dark night. Silent, timeless, jeweled with the constellations of ancient myth and legend, the night sky has inspired wonder throughout the ages."

For most of human history, leadership in studying the heavens has resided elsewhere, but during the 20th century the United States has been the world center of astronomy. This preeminence has been due to good financial support and the imaginative creation of innovative observing equipment. The capabilities of excellent optical telescopes, developed during the first half of this century, were later extended by equipment designed for observing throughout the electromagnetic spectrum. Leading supporters of the development of optical telescopes were the Carnegie Institution of Washington, with its 2.5-meter telescope at Mount Wilson, and the Rockefeller Foundation, which gave the California Institute of Technology funds to build the 5-meter telescope at Mount Palomar. More recently, the National Science Foundation has become a major funder of ground-based astronomy, while NASA has provided excellent facilities in space. The United States has led in exploration of the solar system. In addition, it has launched space vehicles that have permitted observations which could not be achieved from the earth because of absorption of radiation in the atmosphere. The Space Telescope, to be launched in 1985, will be free from atmospheric inhomogeneities that blur sources of light and will be capable of high resolution of objects.

By 1970, generous support of American astronomy had led to many discoveries, including Hubble's expanding universe, time and celestial distance scales, quasars, x-ray sources, high-energy celestial gamma rays, the cosmic microwave background radiation, and polyatomic molecules in interstellar clouds. Discoveries during the 1970's included neutron stars accreting matter from nearby companion stars, hot intergalactic gas whose mass rivals that of the galaxies themselves, vast regions of interstellar gas heated to hundreds of thousands of degrees by shock waves from supernova explosions, and a gravitational lens effect observed as the splitting of light from a distant quasar as the light passed through an intervening galaxy.

---

[*] *Astronomy and Astrophysics for the 1980's*, volume 1, *Report of the Astronomy Survey Committee* (National Academy Press, Washington, D.C., 1982).

The contributions of American astronomy are important and impressive. However, leadership cannot be maintained by resting on our laurels. Continuing preeminence of the United States will be dependent on well-trained people, who are provided with superior equipment.

The astronomical community has made a careful and searching study of opportunities and needs for support for the 1980's. Through extensive consultation and deliberation, a consensus has been achieved. The major new equipment recommended includes (i) an advanced X-ray Astrophysics Facility operated in space; (ii) a Very-Long-Baseline Array of radio telescopes; (iii) a New Technology Telescope, 15 meters in diameter, for ground-based studies in the optical and infrared regions of the spectrum; and (iv) a Large Deployable Reflector in space. All of these proposals would substantially extend the capabilities of astronomy. For example, the Very-Long-Baseline Array would have an angular resolution 100 times better than that of any other image-forming telescope at any wavelength. It would yield detailed radio images of quasars, the nuclei of galaxies, and features of interstellar molecular clouds and other astronomical objects. The first and third items above would be important for many studies, perhaps the most interesting being the examination of extremely distant objects whose radiation was emitted early in the history of the universe.

The report is well constructed and readable. It states well the case for additional expenditures for astronomy. Because of current budgetary problems, its recommendations may not be quickly accepted. However, it is designed to be relevant to the 1980's and at least part of it will surely be ultimately implemented.—Philip H. Abelson

Appendix B

[Reprinted from Astronomy and Astrophysics for the 1980's, Volume 1, Report of the Astronomy Survey Committee (Washington, D.C.: National Academy Press, 1982), pp. 13-22.]

The Astronomy Survey Committee takes note at the outset of the support provided to U.S. astronomy and astrophysics over the past decades through the scientific programs of the National Science Foundation (NSF), the National Aeronautics and Space Administration (NASA), and other federal agencies. This support has enabled U.S. astronomical research to maintain an overall position of world leadership and has vastly widened our horizons for exploration of the Universe.

The programs recommended in this report have been selected from research activities that were, at the beginning of the Survey, candidates for implementation in fiscal year 1983 and beyond. Before presenting a summary of its recommendations, however, the Committee wishes to emphasize the importance of approved, continuing, and previously recommended programs to the progress of astronomical research during the remainder of the decade. The present Committee's recommendations take explicit account of such programs and build upon them.

The Committee calls particular attention to the need for support of the following approved and continuing programs, for which the order of listing carries no implication of priority: Space Telescope and the associated Space Telescope Science Institute, second-generation Space Telescope instrumentation; the Gamma Ray Observatory; NASA level-of-effort observational programs, including research with balloons, aircraft, and sounding rockets, together with the Explorer and Spacelab programs; the Solar Optical Telescope and the Shuttle Infrared Telescope Facility for Spacelab; facilities for the detection of neutrinos from the solar interior; federal grants in support of basic astronomical research at U.S. universities; and programs at the National Astronomy Centers. The 25-Meter Millimeter-Wave Radio Telescope, which was recommended in an earlier form in the Greenstein report, has not yet been implemented. The present status of these approved, continuing, and previously recommended programs is described later in this chapter; their importance for the health of U.S. astronomy in the 1980's is discussed in Chapter 4.

SUMMARY OF THE RECOMMENDED PROGRAM

The Astronomy Survey Committee recommendations for a program in astronomy and astrophysics for the 1980's fall into three general categories:

Prerequisites for new research initiatives;
New programs; and
Programs for study and development.

As noted in the Preface, the observational components of these recommendations are restricted to remote sensing from the Earth or its vicinity. A background and overview of the recommendations follows later in this chapter.

Prerequisites for New Research Initiatives

In order to be effective, the recommended new research initiatives for the coming decade must be supported by a set of Prerequisites that apply to both the gathering and the analysis of the data produced. These Prerequisites are essential for the success of major programs but are inexpensive by comparison. Although significant support already exists for each, the Committee strongly recommends substantial augmentations in the following areas, in which the order of listing carries no implication of priority:

A. *Instrumentation and detectors*, to utilize the latest technology to enhance the efficiency of both new and existing telescopes in the most cost-effective manner;

B. *Theory and data analysis*, to facilitate the rapid analysis and understanding of observational data;

C. *Computational facilities*, to promote data reduction, image processing, and theoretical calculations;

D. *Laboratory astrophysics*, to furnish the atomic, molecular, and nuclear data essential to the interpretation of nearly all astronomical observations; and

E. *Technical support at ground-based observatories*, to ensure that modern astronomical instrumentation is maintained in the best condition permitted by the state of the art.

A detailed consideration and justification of these Research Prerequisites appears in Chapter 5.

New Programs

The Astronomy Survey Committee recommends the approval and funding of new programs in astronomy and astrophysics for the 1980's. These have been arranged into three categories according to the scale of resources required.

A. *Major New Programs* The Committee believes that four major programs are critically important for the rapid and effective progress of astronomical research in the 1980's and is unanimous in recommending the following order of priority:

1. *An Advanced X-Ray Astrophysics Facility* (AXAF) operated as a permanent national observatory in space, to provide x-ray pictures of the Universe comparable in depth and detail with those of the most advanced optical and radio telescopes. Continuing the remarkable development of x-ray technology applied to astronomy during the 1970's, this facility will combine greatly improved angular and spectral resolution with a sensitivity up to one hundred times greater than that of any previous x-ray mission.
2. *A Very-Long-Baseline* (VLB) *Array* of radio telescopes designed to produce radio images with an angular resolution of 0.3 milliarcsecond. Among many potential applications of profound importance, this instrument will probe the small-scale structure surrounding the enigmatic energy sources in the cores of quasars and active galactic nuclei and will directly determine the distance scale within our Galaxy with unprecedented accuracy.
3. *A New Technology Telescope* (NTT) of the 15-m class operating from the ground at wavelengths of 0.3 to 20 μm, to provide a tenfold increase in light-gathering capacity at visual wavelengths and a hundredfold increase in speed for spectroscopy at infrared wavelengths, with application to a very wide range of scientific problems. The Committee finds the scientific merit of this instrument to be as high as that of any other facility considered and emphasizes that its priority ranking does not reflect its scientific importance but rather its state of technological readiness. *The design studies needed before NTT can be constructed are of the highest priority and should be undertaken immediately.*
4. *A Large Deployable Reflector* in space, to carry out spectroscopic and imaging observations in the far-infrared and submillimeter wavelength regions of the spectrum that are inaccessible to study from the ground, thus extending the powerful capabilities of NTT to these longer wavelengths. Such an instrument, in the 10-m class, will present unprecedented opportunities for studying molecular and atomic processes that accompany the formation of stars and planetary systems.

B. *Moderate New Programs* In rough order of priority, these are:

1. *An augmentation to the NASA Explorer program,* which remains a flexible and highly cost-effective means to pursue important new space-science opportunities covering a wide range of objects and nearly every region of the electromagnetic spectrum.

2. *A far-ultraviolet spectrograph in space,* to carry out a thorough study of the 900–1200-Å region of the spectrum, important for studies of stellar evolution, the interstellar medium, and planetary atmospheres.
3. *A space* VLB *interferometry antenna* in low-Earth orbit, to extend the powerful VLBI technique into space in parallel with the rapid completion of a ground-based VLB Array, in order to provide more detailed radio maps of complex sources, greater sky coverage, and higher time resolution than the Array can provide alone.
4. *The construction of optical/infrared telescopes in the 2–5-m class,* to observe transient phenomena, conduct long-term survey and surveillance programs, provide crucially needed ground-based support to space astronomy, and permit the development of instrumentation under realistic observing conditions. The Committee particularly encourages federal assistance for those projects that will also receive significant nonfederal funding for construction and operation.
5. *An Advanced Solar Observatory* in space, to provide observations of our Sun—the nearest star—simultaneously at optical, extreme ultraviolet, gamma-ray, and x-ray wavelengths, to carry out long-term studies of large-scale circulation, internal dynamics, high-energy transient phenomena, and coronal evolution.
6. *A series of cosmic-ray experiments* in space, to promote the study of solar and stellar activity, the interstellar medium, the origin of the elements, and violent solar and cosmic processes.
7. *An astronomical Search for Extraterrestrial Intelligence* (SETI), supported at a modest level, undertaken as a long-term effort rather than as a short-term project, and open to the participation of the general scientific community.

C. *Small New Programs* The program of highest priority is:

- An antenna approximately 10 m in diameter for submillimeter-wave observations, at an excellent ground-based site.

Other programs of outstanding scientific merit, in which the order of listing carries no implication of priority, are as follows:

- A spatial interferometer for observations of high angular resolution in the mid-infrared region of the spectrum;
- A program of high-precision optical astrometry; and
- A temporary program to maintain scientific expertise at U.S. universities during the 1980's through a series of competitive awards to young astronomers.

Detailed discussion and justification of the New Programs appears in Chapter 6.

Programs for Study and Development

Planning and development are often time-consuming, especially for large projects. It is therefore important during the coming decade to begin study and development of programs that appear to have exceptional promise for the 1990's and beyond. Projects and study areas recommended by the Committee in this category include the following, in which the order of listing carries no implication of priority:

A. Future x-ray observatories in space;
B. Instruments for the detection of gravitational waves from astronomical objects;
C. Long-duration spaceflights of infrared telescopes cooled to cryogenic temperatures;
D. A very large telescope in space for optical, ultraviolet, and near-infrared observations;
E. A program of advanced interferometry in the radio, infrared, and optical spectral regions;
F. Advanced gamma-ray experiments; and
G. Astronomical observatories on the Moon.

Detailed discussion of the Programs for Study and Development appears in Chapter 7.

ESTIMATED COST OF THE RECOMMENDATIONS

In order to establish the overall scale of the recommended total program, the Committee gives in Table 2.1 its own approximate estimates of the requirements for new funding over the next 10 years in millions of 1980 dollars. Funds for projects to be supported by NASA represent research-and-development funds within NASA's Office of Space Science and Applications (OSSA); funds for projects to be supported by NSF represent total cost to NSF. Operating costs are included for those facilities expected to become operational in the 1980's.

The funding entries for the Prerequisites for New Research Initiatives represent augmentations to the present levels of support for these activities within NSF and NASA. As it is expected that the two agencies will work together to coordinate support for the Prerequisites, specific agency responsibility is not indicated in the following table. However, since the Prerequisites provide support to space- and ground-based research at comparable levels, the Committee anticipates that the funding augmentations to be provided by NASA and NSF will be roughly equal in magnitude.

In the cases of the New Programs, the division between space- and ground-based projects is clear. Funds listed for the Explorer program represent an augmentation to NASA's level-of-effort budget for that program; the operations costs listed for ground-based projects, together with the temporary program to maintain scientific expertise at U.S. universities, represent further augmentations to the operations budget of NSF's Astronomy Division. Remaining New

TABLE 2.1 Requirements for New Funding

	(Millions of 1980 Dollars)
PREREQUISITES FOR NEW RESEARCH INITIATIVES	
A. Instrumentation and detectors (doubling of present $15 million/year level of effort by increments over 10 years)	$ 75
B. Theory and data analysis (augmentation by $5 million/year)	50
C. Computational facilities (30 minicomputer systems installed at a rate of 5 systems/year, including operations)	20
D. Laboratory astrophysics (augmentation by $2.5 million/year)	25
E. Technical support at ground-based observatories, including 40 new support positions	20
DECADE TOTAL, PREREQUISITES	$190
NEW PROGRAMS	
A. *Major New Programs* In order of priority:	
1. Advanced X-Ray Astrophysics Facility (AXAF)	$500
2. Very-Long-Baseline (VLB) Array (including $15 million for operations)	50
3. New Technology Telescope (NTT)	100
4. Large Deployable Reflector in space	300
Decade Subtotal	$950
B. *Moderate New Programs* In rough order of priority:	
1. Augmentation to Explorer satellite program	$200
2. Far-ultraviolet spectrograph in space	150
3. Space VLB interferometry antenna	60
4. Optical/infrared telescopes in the 2–5-m class	20
5. Advanced Solar Observatory in space	200
6. Cosmic-ray experiments	100
7. An astronomical Search for Extraterrestrial Intelligence (SETI)	20
Decade Subtotal	$750
C. *Small New Programs* Of highest priority:	
• 10-m submillimeter-wave radio antenna (including $2 million for operations)	$ 4
Other important programs:	
• Spatial interferometer for the mid-infrared (including $1 million for operations)	3
• High-precision optical astrometry program	3
• Temporary program to maintain scientific expertise at U.S. universities	10
Decade Subtotal	$ 20
DECADE TOTAL, NEW PROGRAMS	$1,720

Program costs represent new-funding requirements for either NASA (new starts within OSSA) or NSF (major construction within the Astronomy Division).

The cost estimates for AXAF and the VLB Array were derived with the help of individual scientists participating in current studies and are based on reasonably complete rough designs. The actual cost of NTT, however, cannot be estimated until further studies indicate which of several alternative conceptual designs will be most cost-effective; the figure given in Table 2.1 is meant as a limit that the Committee recommends should not be substantially exceeded. The estimated cost of the Large Deployable Reflector in space is highly uncertain because instrumentation of this type has not yet been developed and launched. Most of the costs estimated for the Moderate New Programs should be reasonable approximations, as they are based on experience with previous instruments of a similar nature. The costs given for the augmentation to the NASA Explorer program and for SETI, however, should be regarded as target figures for the level of effort the Committee finds appropriate.

The total cost in new funding estimated for the Prerequisites and New Programs together is about $1.9 billion in 1980 dollars. By comparison, the Greenstein report (1972) recommended new programs with an estimated cost of $844 million in 1970 dollars, or approximately $1.7 billion in 1980 dollars, and most of those programs were in fact implemented. *The program recommended here for the 1980's is thus roughly comparable in scale with that actually carried out during the 1970's on the basis of the recommendations of the Greenstein report.*

The Committee wishes to emphasize, however, that the present recommendations will require substantial increases in the budget of the Astronomy Division of NSF, the agency primarily responsible for the support of ground-based astronomy. If, as anticipated, NSF will provide roughly half of the additional funds required for the Prerequisites for New Research Initiatives, an increase of about 30 percent in the Astronomy Division's operations budget over the real level of expenditures during the 1970's will be required for NSF to carry out its share of the recommended program over the next decade. Funds needed by NSF for major construction over the next 10 years will also be substantially higher than those expended during the 1970's, which saw the completion of only one major project, the Very Large Array, at a cost of $78 million. *The Astronomy Survey Committee believes that these increases in the NSF budget for ground-based astronomy are essential to maintain an effective partnership with space astronomy during the 1980's.*

BACKGROUND AND OVERVIEW

The Greenstein Report

The publication of *Astronomy and Astrophysics for the 1970's* (the Greenstein report) by the National Academy of Sciences in 1972 had a powerful impact on the development of U.S. astronomy and astrophysics during 1972–1982. The federal government on the whole responded positively to its recommendations, with the result that the facilities available to U.S. astronomers have enabled them to remain at the frontiers of research. Here we review the responses to the recommendations of that report and their impact on the progress of science.

Radio Astronomy and the VLA The highest-priority recommendation of the Greenstein report was the construction of a Very Large Array (VLA) radio telescope, together with increased support for smaller facilities. Funded by NSF, the VLA was constructed in stages during the 1970's and was formally dedicated in 1980; by far the largest and most complex ground-based astronomical facility established to date, the VLA was completed on schedule and within budget. VLA studies of radio sources are already having a large impact on both Galactic and extragalactic astronomy (see, for example, the cover of this report).

The recommended increase in funding for smaller radio-astronomical facilities did not materialize, however, nor has funding yet been provided for a recommended millimeter-wave radio telescope, then projected to have a diameter of 65 m and to be operable at wavelengths down to 3 mm. Since the publication of the Greenstein report, the study of interstellar molecules at millimeter wavelengths has yielded insight into the process of star formation; as the science has progressively moved to shorter wavelengths, there has now evolved a need for a smaller, more precisely figured telescope of 25-m diameter, still offering high sensitivity and spatial resolution but operable at wavelengths down to 1 mm. The recommendation for a large centimeter-wave antenna was not implemented, although existing facilities for observations at wavelengths longer than 1 cm have been maintained and in some cases upgraded.

Optical Ground-Based Astronomy The second-priority recommendation was for a variety of steps to enhance the capability available to U.S. optical astronomers. A key proposal was the development and construction of a multiple-mirror telescope (MMT) with aperture equivalent to that of a conventional telescope in the 3.8–5.0-m range; this project was to be followed by the construction of a larger MMT, of 10–15-m aperture, if that proved feasible, or by a conventional telescope of 5-m aperture if it did not. A 4.5-m MMT has in fact been developed jointly by the Smithsonian Institution and the University of Arizona, becoming operational in 1979. However, neither a larger MMT nor a conventional 5-m telescope has been funded, with the result that the largest instrument available to U.S. optical astronomers is still the 5-m Hale telescope on Mt. Palomar, which went into operation 35 years ago. As a result, optical spectroscopy of the faintest galaxies and quasars discovered by radio and x-ray astronomers has not kept pace with new discoveries, even though these extremely distant objects are of great interest because of their bearing on the nature of cosmic evolutionary processes early in the history of the Universe.

The recommendation also called for equipping existing telescopes with advanced electronic detectors and controls. Although major progress was made in the development of such devices during the 1970's, they have so far been provided to only a few major observatories. The capability of instruments now available for use on most of the nation's optical telescopes still lags far behind state-of-the-art technology.

The Greenstein report furthermore recommended that three telescopes in the 2.5-m class be constructed for a variety of purposes. As none of these has been funded, all the nation's major optical telescopes are heavily oversubscribed.

Infrared Astronomy The third-priority recommendation called for an across-the-board increase in support for infrared astronomy, which at that time was beginning to demonstrate its great importance. Support has increased substantially, through the funding of two major ground-based infrared telescopes (the 2.3-m University of Wyoming Infrared Observatory reflector and the 3-m Infrared Telescope Facility operated by NASA on Mauna Kea), the Kuiper Airborne Observatory program, and a balloon program. An international Infrared Astronomy Satellite (IRAS), scheduled for launch in 1982, will carry out a comprehensive, far-infrared survey of the sky, as called for in the Greenstein report.

Appendix C

[Reprinted from Astronomy and Astrophysics for the 1980's, Volume 1, Report of the Astronomy Survey Committee (Washington, D.C.: National Academy Press, 1982), pp. 112-113.]

The Shuttle Infrared Telescope Facility (SIRTF)

The proposed SIRTF will be the cornerstone of research in infrared astronomy during the 1980's. The Astronomy Survey Committee joins with the Space Science Board's Committee on Space Astronomy And Astrophysics (*A Strategy for Space Astronomy and Astrophysics for the 1980's*, National Academy of Sciences, Washington, D.C., 1979) in recommending this facility as the first major infrared telescope in space.

SIRTF will permit investigations over the enormous range of wavelengths from 2 to 300 μm. For some important types of observations it will, because of its cryogenically cooled optics, yield a sensitivity gain of 1000 over the largest existing ground-based and airborne infrared telescopes; this gain in sensitivity is so large that it is not unreasonable to expect that SIRTF will make important and unexpected discoveries. The multiple, interchangeable focal-plane instruments planned for SIRTF will moreover greatly increase our ability to explore the evolution of distant extragalactic sources, the physical properties and chemical composition of molecular clouds and regions of star formation, the nature of cometary nuclei and asteroids, and the structure of planetary atmospheres. For example, SIRTF will be able to detect infrared sources at the limits of the observable Universe; on one Shuttle flight, it could gather information on sources of both large and small red shift, thus permitting a comparison of the energetics of quasars and galaxies at the earliest epochs of the Universe with those at the present epoch. Because of its relatively wide field of view, SIRTF will also be able to carry out efficient surveys of infrared sources that will help to optimize the observing programs of larger instruments, such as the New Technology Telescope (NTT), the VLB Array, and a Large Deployable Reflector (LDR) in space, which have narrower fields of view.

SIRTF should be an early and frequently flown payload on Shuttle sortie missions. It has such high sensitivity that very extensive infrared observations can be accomplished even within the relatively brief, 7-day profile of an early Shuttle flight; reflight on the Shuttle makes it possible to fly SIRTF with continually improved focal-plane instrumentation and detectors. Eventually, however, SIRTF flights of longer duration will be needed to realize the full potential of this remarkable facility, and the study of such flights is recommended in Chapter 7.

Appendix D

[Reprinted from Astronomy and Astrophysics for the 1980's, Volume 1, Report of the Astronomy Survey Committee (Washington, D.C.: National Academy Press, 1982), pp. 135-136.]

2. *A Very-Long-Baseline* (VLB) *Array of Radio Telescopes*

The Astronomy Survey Committee recommends the construction of a ground-based Very-Long-Baseline (VLB) Array of radio telescopes designed to produce images with an angular resolution of 0.3 milliarcsecond. Because the Array utilizes proven technology, this project may be begun immediately after completion of final management and design studies.

Extraordinarily high angular resolution is now possible at radio frequencies. Precision atomic clocks, more sensitive and reliable receivers, high-speed tape recorders, sophisticated image-processing techniques, and modern antennas now make it feasible to build a radio array with the angular resolution of a telescope covering an entire continent. This may be done by synchronizing the operation of about ten widely spaced antennas of approximately 25-m diameter, whose outputs are recorded and later combined in a central computer.

This VLB Array will produce high-quality radio images capable of resolving features down to 0.3 milliarcsecond (the size of a dime in New York as seen from Los Angeles). This is a hundred times better angular resolution than that of any other image-forming telescope at any wavelength and will yield detailed new radio images of a wide range of astronomical objects at the frontiers of modern astrophysical research. These include quasars and the nuclei of galaxies, features of interstellar molecular clouds, the center of our Galaxy, and a variety of energetic Galactic objects such as x-ray, binary, and flare stars. The high angular resolution of the VLB Array will permit the direct study of small-scale structure surrounding the central regions of quasars and stars in the process of formation. Through the method of statistical parallaxes, it will furthermore permit direct measurements of distances to many objects throughout our Galaxy and even to some in nearby galaxies. The VLB Array can also be applied to important problems in Earth science (including precision geodesy and geophysics), to the navigation of interplanetary spacecraft, and to tests of the General Theory of Relativity.

Although the VLB Array is a complex and sophisticated instrument, it will make use of proven concepts and instrumentation. Construction should begin immediately upon completion of management and design studies with the building of the antennas and the development of the data-reduction system and other instrumentation. Collaboration with groups in other countries, particularly in Europe and North America, would improve the performance of the instrument by increasing the resolution even further (particularly in the north-south direction) and by improving the image quality at low declinations.

REPORT OF THE BRIEFING PANEL ON AGRICULTURAL RESEARCH

Briefing Panel on Agricultural Research

Lawrence Bogorad (Cochairman), Harvard University
William L. Brown (Cochairman), Pioneer Hi-Bred International, Inc.
Myron K. Brakke, USDA-ARS, University of Nebraska
Vincent G. Dethier, University of Massachusetts
Michael Freeling, University of California, Berkeley
Ralph W. F. Hardy, E.I. duPont de Nemours & Company, Inc.
Gary H. Heichel, USDA-ARS, University of Minnesota
Arthur Kelman, University of Wisconsin
David W. Krogmann, Purdue University
Carole P. Meredith, University of California, Davis
Oliver E. Nelson, Jr., University of Wisconsin

Staff

James E. Tavares, Senior Staff Officer, Board on Agriculture
Allan R. Hoffman, Executive Director, Committee on Science, Engineering, and Public Policy

REPORT OF THE BRIEFING PANEL ON AGRICULTURAL RESEARCH

SUMMARY

This report focuses almost exclusively on plant science research related to improved agricultural production efficiency. An assessment is made of the realistic expectations of genetic engineering applications for improving crops and cropping practices; several promising research areas for increased federal research investment are identified; and some constraints to plant science research are briefly discussed.

Genetic Engineering

In the near future the biggest contribution of the application of plant molecular biology and genetic engineering techniques to plants will be to fundamental knowledge. Such new knowledge in plant science will have enormous potential for making major contributions to crop improvement and cropping practices.

Genetic engineering has already provided unexpected new insights into the organization and function of genetic traits. The direct demonstration of relationships between regulatory and structural genes, the presence of repeated sequences of DNA throughout the genome, the existence of introns or noncoding sequences within genes, and the surprising plasticity of the genome that was earlier thought to be rigidly conservative have all helped and forced scientists to think in new ways about the factors that influence and control the reassortment and expression of hereditary traits. Much of this increased understanding could only be derived through the use and analysis of the technologies associated with genetic engineering.

Genetic engineering is a powerful adjunct to plant breeding. The successful application of genetic engineering will generate new and previously unavailable sources of desirable genetic variability that can be exploited through conventional plant-breeding techniques. The application of cell culture techniques is already contributing new sources of genetic variation to plant breeding.

While the application of gene-splicing techniques in plants is still in its infancy, it seems reasonable to expect that there are traits controlled by a single gene or a small group of genes (such as for herbicide resistance) that can be transferred successfully using genetic engineering. Important questions remain, however, on whether the introduction of such foreign DNA will affect the expression of other genes in the plant and whether coordinated multigene traits can be engineered successfully.

Research Opportunities

Advances in molecular biology techniques and improved instrumentation are providing new opportunities in plant science research. In the mid-1970s several high-priority research areas in the plant sciences were identified, and four of these areas--photosynthetic enhancement, biological nitrogen fixation, genetic mechanisms, and plant biological stress--received further recognition through increases in federal financial support. The resulting expansion of research programs has been highly productive.

The following seven research opportunity areas were identified by the Briefing Panel on Agricultural Research as holding particular promise for establishing the basic scientific understanding that could lead to important discoveries and applications for crop improvement. These areas are listed in order of priority.

Genome Structure and Gene Expression in Plants

Basic to all plant science research is an understanding of the genetic information system that controls plant growth and function. Essential to progress in these areas are the identification and isolation of specific genes in plants, the development of techniques for inserting the DNA of other plants and organisms into plant cells, and the establishment of *in vitro* systems that will permit the genetic expression of such foreign DNA.

Plant Developmental Biology

Discoveries in the regulation of gene expression in plants will greatly aid in an understanding of the control and development of form and function in plants. Conversely, the elucidation of factors involved in intercellular communication and the coordination of processes on an organismic level in plants should provide important insights for studying the genetic controls of plant development.

Plant and Pest Interactions

The genetic and biochemical mechanisms that control virulence in plant pathogens and the related resistance that can be established in crop cultivars are now amenable to analysis because of the availability of the powerful tools of molecular biology. The wealth of information on genome structure and the control of gene expression in viruses and bacteria can hasten advances in the understanding of the actions of these and other pathogens. Continued and accelerated work in identifying bacterial toxins and host-specific fungal toxins is needed, along with work on the action of chemicals that trigger systemic defense mechanisms in host plants. Multidisciplinary studies combining insect neurobiology and the analytical chemistry of secondary plant metabolites that attract or repel insects should lead to useful applications for controlling pest insects.

Organisms in the Rhizosphere

Little is known about the ecology of the organisms that inhabit the root zone. The intriguing properties of disease-suppressive soils and the growing recognition of the importance of the ubiquitous mycorrhizal fungi point to the enormous potential benefits that could be derived from intensive study of these soil organisms.

Photosynthetic Energy Conversion and Carbon Metabolism

This report discusses in some detail the considerable progress that has been made in understanding the components of photosynthetic energy conversion. The methods and research tools are at hand to continue and expand this progress. In addition, increased emphasis needs to be placed on research in carbon metabolism, particularly on the control of partitioning the carbohydrate products of photosynthesis throughout the plant and on the feedback effects of this partitioning on net photosynthesis.

Mineral Uptake, Deficiencies, and Excesses

The existence of crop cultivars with impaired or improved uptake of essential mineral nutrients or resistance to toxic levels of other minerals provides opportunities for applying the tools of molecular biology to better understand mineral nutrition in plants. The benefits to be derived are twofold, namely, the more efficient production of crops on soils currently supplemented with high-nutrient inputs (particularly nitrogen) via a reduction in inputs, and the extension of cropping to marginal soils that are saline or contain toxic minerals.

Biological Nitrogen Fixation

Advances in the understanding of biological nitrogen fixation continue at all levels of research, from sequencing the nucleic acids in the genes of nitrogen-fixation systems to practical applications such as selecting or devising improved strains of nitrogen-fixing bacteria in their associations with crop plants. Many intriguing problems remain, such as why legumes do not show any yield response to nitrogen fertilizer as other crops do, how to control the wasteful loss of reducing power in the form of hydrogen gas generated by the nitrogen-fixation system, and how to achieve the ultimate goal of transferring functioning nitrogen-fixation genes directly into crop plants.

Research Constraints

To realize the full potential of the research opportunities identified in this report, some attention also should be paid to factors, in addition to adequate funding, that influence the rate of progress in plant science research. These factors include the need for trained scientists and for institutional arrangements to encourage multidisciplinary research and to foster innovative scouting research.

Currently there is a shortage of scientists with training in both the basic plant sciences and agronomy as well as molecular biology and molecular genetics. This deficiency might be addressed in part through postdoctoral research opportunities to permit recent graduates in either plant science fields or molecular genetics to obtain research experience in the other area.. At present there are relatively few postdoctorals in the agricultural sciences. Recent data published by the National Science Foundation show that there were 23,420 postdoctorals supported by federal sources in biological sciences, and 15,845 in health sciences, but only 821 postdoctorals in agricultural sciences for the period 1974 through 1979.* An alternative method for stimulating the interchange of ideas among disciplines and fostering cooperative research might be through special summer study programs. Two examples that might serve as models are the Cold Spring Harbor schools in molecular biology and molecular genetics and the Woods Hole Marine Biological Laboratory summer programs in several areas of biology. Both examples have been instrumental in advancing research in their areas.

There are perhaps additional ideas on how interdisciplinary research might be fostered, including the encouragement of multidisciplinary research teams and the opportunities to support the exploratory or scouting research that provides the foundation on which new, focused research programs and potential technologies are built.

INTRODUCTION

U.S. agriculture is continually changing. In contrast to the kinds of changes that occurred in earlier periods of the nation's history, the pace and nature of changes in agriculture today are to a considerable extent the result of scientific research and development. Such changes have helped account for this country's ability to supply itself with abundant, high-quality food and fiber at low cost and to help meet some of the food needs of the rest of the world.

---

*National Science Foundation. 1981. National Patterns of Science and Technology Resources. NSF 81-131 (Washington, D.C.: U.S. Government Printing Office).

Notable examples of wide-ranging effects of past research that have helped secure the success of the American agricultural system include the following. Fertilizers now permit the bountiful harvests that virgin soils were once capable of producing. Agricultural chemicals protect crops from insects and pathogens and have largely replaced mechanical cultivation practices that once controlled weeds. These same chemicals have found new uses in no-till or minimum-till farming that helps reduce one of the major problems of modern agriculture--soil erosion. Modern plant-breeding technology, based on a strong foundation of plant genetics, has produced crop varieties that are well adapted to particular environments. Plant breeders have kept ahead of continuously evolving pest populations with a steady stream of resistant varieties.

Obviously, as with any type of scientific progress, these results in American agriculture were not brought about overnight by single-handed effort or by a few isolated discoveries. The results have come about in the same way that any organism grows--through a vast number of interrelated developments. The sustained success of American agriculture depends upon the continuation of research programs attuned to the problems of the present and to the changes needed in the future.

This report focuses almost exclusively on plant science research, the chief beneficiary of which traditionally has been agriculture. The report summarizes the discussions of a panel of 11 members who were asked to identify promising research opportunities in the plant sciences that could lead to developing more productive disease-resistant, stress-tolerant, and energy-efficient crop plants and to provide some assessment of the realistic expectations of genetic engineering for developing such crops. During three days of discussion several areas were identified as being particularly promising opportunities for highly effective use of incremental increases in federal research investment. These research areas include the following: genome structure and gene expression in plants; plant developmental biology; plant and pest interactions; organisms in the rhizosphere; photosynthetic energy conversion and carbon metabolism; mineral uptake, deficiencies, and excesses; and biological nitrogen fixation.

RESEARCH OPPORTUNITIES

Genome Structure and Gene Expression in Plants

In a crop plant there are approximately 100,000 genes in each nuclear genome, that is, the complete set of genes found in the nucleus of each plant cell. At its simplest level, each gene is a DNA (deoxyribonucleic acid) sequence designed along similar lines, and each involves a unit of information used to specify a product called messenger RNA (ribonucleic acid), which in turn dictates the structure of proteins that control the plant's functions. Little is known about the molecular structure and organization of the genome or about factors regulating the expression of the individual genes. This is true of animals as well as of plants. An understanding of genome structure and gene function in plants is growing, however, because the means of acquiring answers to many questions are now available. The new tools of molecular biology and genetic engineering that offer prospects of identifying, isolating, and cloning genes will permit an understanding of the control of gene expression with a view to transferring such genes into crop plants.

The generalized function of some of the DNA sequences within a gene is known. There are DNA coding sequences that specify the amino acid sequence of proteins, and DNA promoter sequences that may bind enzyme molecules that initiate or prevent the initiation of the transcription of the gene into RNA. Stretches of DNA, called introns, reside within the coding sequences of most nuclear genes and are present in the RNA first transcribed from the DNA sequence, but are removed from this primary RNA as it becomes the messenger RNA that is translated into protein. The function of introns is not known. There are DNA sequences within or next to coding sequences that are thought to bind molecules that regulate the tissue-specific timing of gene expression during development. Also, there is evidence that changes in the organization or arrangement of genes in the genome can influence gene function. These levels of gene regulation are not understood. Fewer than 10 plant genes have been fully described as DNA sequences.

Recent advances in molecular biology include recombinant-DNA technology and rapid DNA-sequencing methods. In addition, transposable elements--DNA sequences that have the ability to move from place to place within the genome and affect the functioning of genes into which they insert--can be used as intragenomic probes to identify and isolate genes they associate with and are among the powerful techniques now available for gaining an understanding of the structure and regulation of genes in plants. Transposons, comparable mobile DNA sequences in bacteria, have been used successfully as such probes.

Specific areas requiring research include identification of which genes code for particular enzymes in plants and what molecules elicit the activity of such genes in a certain tissue or at a particular stage of development. An important extension of this knowledge would be the determination of what gene or genes could be transferred from one species to another with a view to improving crop productivity. This determination also requires the identification and understanding in molecular terms of the rate-limiting steps in plant growth and metabolism.

Knowledge of specific gene expression systems may provide information leading to control of rate-limiting steps. An example is the regulation of the bacterial nitrogen-fixation gene system by the chemical methionine sulfoxamine (MS). In the presence of high levels of ammonia or fixed nitrogen, bacterial nitrogen fixation will not occur. However, added MS causes the nitrogen-fixation gene system to be expressed even in the presence of quantities of fixed nitrogen that otherwise prevent expression of the system. This effect of MS provides an important example of how an understanding of the mechanism that regulates the expression of a biosynthetic system might lead to the identification and synthesis of specific growth-regulatory products that promote or arrest the system.

Plant growth regulators for increasing yield and improving crop quality have been sought for some time, but the success in identifying regulators for major agronomic crops has been very low. An understanding of plant gene structure and of gene expression systems should provide the essential knowledge base for identifying and synthesizing new plant growth regulators. Many suggest that the benefits of such as-yet-undiscovered plant growth regulators to agriculture may equal or exceed the benefits of plant protectants.

A particularly important area within the field of gene organization and expression involves the development of systems in which cloned DNA sequences may be assayed for function. The primary function of the DNA coding sequence of the gene involves the formation of an RNA intermediate, but this function for many genes is evoked only at particular times in development and often only in particular cells of the organism. There is specific need at present for the development of additional *in vitro* systems that permit the transcription of DNA coding sequences to RNA and their further translation to protein-enzyme products. More importantly, although astounding progress has been made in developing the crown gall bacterial plasmid system for plant genetic engineering, there is need for additional *in vivo* transformation systems in which genes can be integrated into the chromosomes of a cell, and that cell can, in turn, be used to generate an entire plant within which the gene's expression may be assayed. The cell wall can be removed from certain kinds of plant cells to produce protoplasts, and in some species these protoplasts can be regenerated into whole plants. Current research involving the introduction of foreign DNA into protoplasts is particularly promising.

In addition to the 100,000 or so nuclear genes in each plant cell, two plant cell organelles also contain genes. The cell's chloroplasts--the seat of photosynthesis--each contain perhaps 100 unique genes; and mitochondria--the sites for oxidation of foodstuffs in the cell--also contain many genes. These three genetic systems of plant cells are highly interdependent, although features of their gene expression systems differ. The health, vigor, and success of a crop variety depend on organelle as well as nuclear genomes and on the proper integration of all three genetic systems. Furthermore, each of these genomes may have its own advantages as a site for introduction of genes for crop improvement.

A few laboratories in the United States currently are intensively investigating DNA sequences and their functions in plants. The knowledge base that these laboratories are developing is essential to the application of this technology to crop improvement.

Plant Developmental Biology

Plant developmental biology is the study of how a single cell gives rise to an entire plant with all its different tissues, organs, and functions. Much of the work in this area has been descriptive, but with the emergence of more sophisticated techniques, an understanding of molecular mechanisms of plant development is now possible. New and powerful tools of cellular and molecular biology promise to provide the means for achieving a much deeper understanding of the ultimate basis for the development of form and function in plants.

The identification and study of those genes and regulatory elements in the plant genome that have profound effects on development will lead to important advances in current knowledge of how the many genes in each cell are coordinately expressed to produce a complex organism with many integrated functions. As these important genes and sequences are identified, they can be modified, rearranged, or otherwise manipulated, and the effects on development can be analyzed. The study of unique kinds of mutants (called homeotic mutants) in the fruit fly, *Drosophila melanogaster*, have been invaluable in revealing how major developmental processes in this organism are controlled. Similar activity in plant developmental biology could produce comparably significant breakthroughs in the understanding of plant growth.

Plant cell culture is an additional tool with potential for facilitating advances in plant developmental biology. Some developmental processes can be controlled in cell cultures; a notable example of such a process is somatic embryogenesis, the development of embryos, and then plants, from cells in culture. This high degree of control and the capacity for manipulation that plant cell culture provides permit a developmental process to be intensively and effectively studied in an appropriate cell culture system. For example, mutant carrot cell cultures that are blocked at different stages of somatic embryogenesis are currently facilitating the greater understanding of this developmental process.

Further research in plant developmental biology will result in greater understanding of the crop plant as a biological organism. A number of specific areas can benefit directly from increased knowledge in this field. Yield is the sum of a number of components, many of which are based in plant form and function. The shape of a plant may critically influence its productivity. For example, new wheat and rice varieties with shorter stalks are already improving yields in many parts of the world. Branching patterns and leaf morphology can influence the severity of disease. Further understanding of the biological basis for plant form might permit more extensive manipulation of form with a resulting change in productivity. There is some information on the role that several of the recognized plant hormones play in communication between plant cells and between different parts of a plant. Additional understanding of the coordination of processes on an organismic level might also inspire the development of new plant growth regulators for the deliberate modification of agriculturally important processes such as senescence and assimilate partitioning (the distribution of photosynthetic assimilates to different parts of a plant). Finally, there is no question that the regeneration of plants from cell cultures of important crop species such as grain legumes and cereals will be improved by further knowledge and will thus facilitate genetic engineering with these crops.

Plant and Pest Interactions

Although dramatic successes in disease control and to a lesser extent in insect control have been achieved through a variety of plant-breeding techniques, the basic genetic, biochemical, and physiological mechanisms that enable plants to resist attack by pathogens and insects remain largely unknown. Effective chemical controls for diseases of plants caused by bacteria and viruses are lacking at present, and many chemicals used to control fungal, nematode, weed, and insect pests cause concern in terms of their safety to humans and the environment. In addition, pest populations have demonstrated the ability to acquire resistance to pesticides, pointing to the need for greater emphasis on the search for alternative approaches to minimize losses caused by plant pests.

The small size and relative simplicity of the genome in bacteria, viruses, and viroids--the latter being a very small pathogen, smaller than viruses--suggest that available molecular biology and genetic engineering techniques could be used to determine the factors that control virulence in the pathogen and thereby contribute to the development of more effective methods of control. The structure of the simplest pathogens, viruses and viroids, consists of no more than a nucleic acid sequence of DNA or RNA and, in the case of viruses, a protein coat. The key to recognition by a virus of a specific plant host and resistance to viral damage to the host probably will be revealed in the details of the role that the host's nucleic acid and protein-synthesizing systems have in replicating the virus. Viroids may be the first pathogen for which the disease-causing mechanism is known. Their RNAs have been sequenced and are small enough (a few hundred bases) so that specific point mutants can be constructed and the mechanism of their pathogenicity tested. One hypothesis is that viroids may interfere with the normal functioning of a host's small nuclear RNAs. These nuclear RNAs are thought to be involved in removing introns and splicing them out of host RNA transcripts.

Viruses are carried from plant to plant by vectors such as insects, mites, nematodes, or soil fungi. In a few cases, viruses multiply within insect vectors, as well as in the host plant. This remarkable versatility of replication contrasts with the specificity between host and virus illustrated by the failure of the same viruses to multiply in closely related plants or insects. The molecular basis for this highly specific recognition of divergent hosts is unknown, but viral mutants exist that have lost the ability to multiply in one of their hosts but not in others. Such mutants should be useful in investigating this problem.

Recent advances in bacterial genetics provide new tools for studying the underlying genetic and biochemical mechanisms that control virulence factors in bacteria. Knowledge of virulence factors will aid in determining mechanisms of plant resistance to pathogenic bacteria. A lack of understanding of the genetics of the attacking organisms has frustrated previous attempts to analyze the complexities of the biochemical reactions triggered when pathogens invade susceptible or resistant host plants. Most of the bacteria that cause diseases in plants are strains of *Erwinia* and *Pseudomonas*. A wealth of information is now available on the genetics of closely related bacteria such as *E. coli* and other *Pseudomonas* species, and exploitation of this information and use of new techniques are just beginning. Recent advances in studies with the crown gall bacterium *Agrobacterium tumefaciens* and its tumor-inducing (Ti) plasmid have elucidated the molecular biology of the interaction between this plant parasite and its plant hosts. Although crown gall is an atypical plant disease, results obtained using it indicate the possible advances that could be made with other pathogens.

The work with the Ti plasmid has progressed particularly rapidly because of its potential use as a vector for introducing foreign DNA into plant cells. Studies with other bacterial or viral pathogens also could lead to potential vectors for use in genetic engineering.

Certain groups of nonpathogenic bacteria have a remarkable ability to exist and multiply on leaf surfaces. Little is known about the role of such bacteria in aiding or antagonizing invasion of the host by pathogenic bacteria or fungi. It is possible that these bacteria can be used as antagonists of plant-pathogenic forms directly. Strains might be altered genetically to survive and antagonize plant-pathogenic bacteria and fungi.

Progress has been made recently in identifying the chemical structure of bacterial toxins and host-specific fungal toxins. A number of elicitors, compounds that trigger resistance reactions in plants, have also been identified. Key areas for research are those required to determine both the nature of receptors in plants that react to elicitor molecules, which are formed either by the pathogen or indirectly as a result of their invasion of the host, and the nature of the messenger molecules that are transmitted throughout the host plant and initiate the production of compounds involved in the resistance reaction. For example, one such compound, described as a proteolytic enzyme inhibitor-inducing factor, is released by damage to almost any part of a tomato plant; it spreads systematically throughout the plant and triggers the production of a compound that inhibits enzymes that digest proteins. It is likely that other compounds directly or indirectly inhibitory to bacteria and fungi can be formed in plants in a similar manner. Knowledge of such systems can provide clues to identifying resistance mechanisms in plants, can aid in the selection of resistant crop varieties, and can lead to the design of chemicals that elicit the resistance reaction.

In the case of insects, the relationship between host and herbivore is based on a behavioral response on the part of the insect to secondary plant metabolites. To understand these interactions it will be necessary to know the identity of the plant metabolites that stimulate or inhibit feeding by the insect, the characteristics of the insect's chemical sensory system that permit detection of different plant metabolites, and the nature and mechanism of any plant responses to insect attacks. For example, insect neurobiology studies are identifying new neurotransmitters, including some that are apparently unique to insects and may lead to new insect control methods. Also, many insects do the greatest damage while they are feeding at a larval stage. They are maintained in this stage by juvenile hormones. Some plants can produce compounds that interfere with the action of the insect juvenile hormone and result in an early termination of the larval stage, and thus reduce damage to the plant.

New techniques are now available for understanding chemosensory systems (e.g., electrophysiology, biochemistry of receptor proteins, and application of information theory to sensory coding and central nervous system integration) and for identifying plant volatiles, phenolics, alkaloids, and so forth (e.g., gas chromatography and mass spectroscopy). Organized collaborative efforts of plant product chemists, plant physiologists, and electrophysiologists should be aimed at identifying those plant chemicals that attract insects and induce feeding or oviposition and those compounds that prevent these activities by acting as deterrents and as feeding suppressants. Such approaches now hold great promise for the development of improved methods of control and for the selection of resistant plant varieties.

Organisms in the Rhizosphere

Soils are the primary source of nutrients for plants, animals, and people. Extensive research has been done on the physical and chemical nature of soils, but the microbial ecology of the rhizosphere, the soil-root zone, has been neglected even though some of the most destructive diseases of plants are those caused by soil-inhabiting organisms. These diseases are among the most difficult of all to control. Treatment with chemicals is costly, and effective compounds can often cause dramatic shifts in populations of nontarget organisms and may result in shifts to different pathogens, which would be similar to the unanticipated and undesirable shifts between beneficial and harmful insect populations experienced with the use of insecticides.

Major advances in the control of soil-borne pathogens have resulted from the development of resistant cultivars; however, additional progress in reducing crop losses from these pathogens may depend on the use of biological control systems, because effective resistance systems are not available. Some recent studies in the states of California and Washington on disease-suppressive soils indicate that certain strains of the bacterium Pseudomonas have the ability to colonize the surfaces of roots of a number of crop species, resulting in enhancement of root development and increased yields. Applications of such strains of bacteria to wheat seed have resulted in impressive yield increases on soils infested with the destructive fungal root pathogen that causes the "take-all" disease. It is possible that genetic engineering can be used to enhance the effectiveness of root-colonizing organisms as biological control agents.

Another major area for additional study involves the utilization of symbiotic fungi that form intimate associations with roots and enhance nutrient uptake and root development. These associations are called mycorrhizae. Two basic types of mycorrhizal fungus exist: the ectomycorrhizal forms that live on the outer tissues of roots and the endomycorrhizae that invade and live primarily inside root tissues.

Ectomycorrhizae have been shown to be particularly important to forest trees, especially for seedling growth. The potential for large-scale exploitation of mycorrhizae through inoculation of tree seedlings with improved mycorrhizal fungal strains requires additional research on the biology, genetics, and ecology of soil organisms, as well as the study of which trees are physiologically best adapted to mycorrhizal relationships.

Endomycorrhizae are found primarily in the roots of annual and perennial grasses and shrubs, and their value to plant growth has been clearly shown. It has not yet been learned how endomycorrhizal fungi can be cultured and used as inoculants for seeds or seedlings, but advances in the knowledge of how to manipulate these ubiquitous and important fungi have great potential for beneficial effects on crop production.

Minimum- and no-till cropping practices have been used to counteract severe soil erosion conditions and to save on the fuel costs of operating farm machinery, but they leave organic debris on the soil surface, resulting in soil that is wetter and colder. Both factors have important effects on the population dynamics of soil organisms. In addition, the incorporation of fertilizers and organic matter into soil and the aeration and compaction of soil by tillage affect growth conditions for soil microorganisms in the rhizosphere and, in turn, crop productivity and efficiency. For example, as some microorganisms degrade surface debris they produce toxins that may inhibit seedling growth in the crop that follows. Also, evidence suggests that intensive tillage and heavy fertilizer applications have helped counteract suppression of crop growth caused by a variety of universal, debilitating but nonfatal root diseases. If beneficial bacteria such as the Pseudomonas strains noted above are to be utilized to reduce the expense of mechanical tillage and fertilizers, a better understanding of the ecology of the rhizosphere and interactions among soil organisms is needed.

The evidence of potential, beneficial effects from symbiotic mycorrhizal fungi and certain other beneficial root-colonizing bacteria and fungi highlights the need for a better understanding of the ecology and population dynamics of soil microorganisms and the effects of different cropping practices on these microorganisms. Some important areas of research include the role of soil microorganisms in nutrient cycling as they affect the availability of mineral nutrients to crops; elucidation of the inherent factors that determine the relationship of bacteria, fungi, and nematodes with plant roots; the genetics of soil microorganisms and the extent of genetic exchange among them; and the biological factors associated with those unique conditions that result in so-called disease-suppressive soils--soils in which pathogens are present but in which little or no crop root damage occurs. Techniques are now available for analyzing the biological and physical factors that would enable scientists to understand the basis for suppression of these pathogens.

Proper recognition of the ecology of soil microorganisms may lead to production of crops with less fertilizer, less soil erosion, and less tillage. Proper understanding of the ecology of the various organisms can result in decreased pest and soil erosion problems, which would permit development of farming systems with increased efficiency in crop production. Such systems might require more crop rotation, multiple cropping, minimum tillage, special green-fertilizer crops, new crops, or varieties bred for a specific cropping system. Basic information is essential for construction of the conceptual models required for rational manipulation of rhizosphere populations to minimize adverse effects of disease agents and to enhance the benefits from symbiotic organisms.

Photosynthetic Energy Conversion and Carbon Metabolism

Photosynthesis is the process on which all agricultural productivity depends. It is unlikely that significant ways for improving photosynthetic efficiency can be expected until the component steps of photosynthesis and their regulation in plants are understood. Understanding of the biological process could also be expected to help improve the efficiency of physical and chemical systems for solar energy conversion.

The first step in photosynthesis is the physical process in which light energy drives the separation of electrical charges. In green plants the charge separation occurs very efficiently, and the subsequent stabilization of the products of this reaction produces very high efficiency in energy conversion. With the availability of spectroscopic techniques to follow the movement of electrons at time resolutions in the subpicosecond range, the biophysical deciphering of the energy conversion process is well under way.

The stabilized photochemical products must be processed into chemical agents that can be utilized in biochemical processes within living cells. This processing occurs in solid-state devices--the photosynthetic membranes assembled within the chloroplast. There has been good progress in taking these membranes apart, purifying individual protein components, and recognizing their three-dimensional structure. So far only a few of the approximately 100 components have been studied completely, but the problem is manageable. The techniques of biochemistry, molecular genetics, and x-ray crystallography are equal to the task of taking the chloroplast apart, sorting, identifying, and understanding the pieces in terms of their individual shapes and functions, and then reassembling the pieces to validate the understanding of the whole.

Specific examples from the research accomplishments of the last few years demonstrate the potential for solving the structure of the photosynthetic apparatus. The enzyme (a carboxylase) that first captures carbon dioxide from the atmosphere to begin its conversion to carbohydrate is an exceedingly large and hard-to-handle aggregate of protein subunits. The smaller protein subunit piece is amenable to an analysis of its amino acid sequence. This has been accomplished by clever but conventional chemical techniques. The larger subunit piece of this enzyme is, by nature of its size, a forbidding task for chemical sequencing. The gene that directs the synthesis of the large subunit has been isolated, and this DNA was inserted into a bacterium so that it could be copied to an amount and condition that made further analysis much simpler. The sequence of nucleotide subunits that made up the genetic instructions that specify the sequence of amino acids in the large protein subunit was determined by an easily used enzymatic technique. By applying previously developed code translation techniques, the DNA sequence of nucleotides then could be translated into an amino acid sequence. This application of genetic engineering techniques simplified--and consequently shortened by many years--the task of determining the sequence of amino acids in the large subunit protein and the subsequent mapping of its enzymatically active sites; furthermore, these techniques can be applied in principle to each of the proteins involved in the photosynthetic process.

The next step needed is to learn the arrangement of this chain of amino acids in the subunits in three-dimensional space. An x-ray crystallographic analysis of the purified protein is under way and will tell the exact details of how the chain is folded and how the subunits are organized in a functional aggregate. Studies now in progress describe how the enzyme works and which parts of it are devoted to catalysis or regulatory modulation. While these studies are currently using chemical probing and alteration of the protein structure, it is clear that genetic manipulation (engineering) will make this process go much more quickly.

Finally, it is necessary to find out where this enzyme sits on the photosynthetic membrane and how it interacts with neighboring proteins that are its reaction partners. Because of its large size, the carboxylase enzyme can be recognized on the surface of the chloroplast membrane by electron microscopy. Certain techniques can be applied to identify neighboring molecules. When this work is complete, this component will be known just as the mainspring in a watch is known. The patient application of these techniques to the other parts of the mechanism is needed.

There is good evidence that many details of the basic process of photosynthesis are identical in all plant species. When the problem has been solved for one, it will have been largely solved for all. Furthermore, the energy processing that occurs in chloroplast membranes is very similar to energy processing in all types of cells. The detailed topographical analysis of the chloroplast membrane should provide powerful insights into energy conversion processes in every living system.

A wasteful reaction associated with photosynthetic reduction of carbon dioxide (CO_2) occurs in many plants. This reaction is known as photorespiration and can account for as much as a 50 percent loss in net photosynthesis. Recent research advances in photorespiration include the genetic verification of the biochemical pathway related to this process. Support of research on photorespiration in recent years has resulted in identifying some of the molecules that might be altered to eliminate or reduce photorespiration losses. Evidence points to the fact that the carboxylase enzyme that converts CO_2 to reduced carbon also carries out photorespiration when oxygen is present.

Research is needed so that photosynthesis can be understood beyond the chloroplast compartment where energy conversion and processing occur. Carbohydrate is the finished product of the chloroplast, and it must be exported to all parts of the plant. It is evident that there are intricate signals that regulate this export process. The filling up of storage reservoirs such as roots, tubers, or seeds with carbohydrates or the deposition of starch in the chloroplast can diminish the rate of leaf photosynthesis. Membrane transport regulators and enzyme modulators are being recognized as some of the signal mechanisms involved. It is likely that these processes can be elucidated and with considerable practical results, such as better controlling the overall efficiency of photosynthesis and guiding the distribution of products into the harvestable parts on the plant.

Understanding the partitioning throughout the plant of the carbohydrate products of photosynthesis should enable a directed approach to increase their movement into the part of the plant harvested. The importance of such an approach is supported by the fact that most of the yield increase in wheat over the last several decades is directly correlated with empirical improvement of assimilate partitioning into the seed.

Assimilate-partitioning research is ready for an expanded effort. Assimilates include sucrose, the major transportable product of photosynthesis, but also include other compounds such as amino acids and other nitrogen-containing molecules. Initial physiological and molecular understanding of the route of the transfer of sucrose includes proton-coupled sucrose cotransport driven by an ATPase to deliver the sucrose to the plant's circulatory system. Similar understanding of the energy-driven transfer in transporting sucrose into the seed has been developed. The compounds of nitrogen that are transported from leaves to the seeds are being explored.

A large, emerging problem area in photosynthesis has to do with understanding the energy costs of biological processes and the relation of these costs to the energy source in the chloroplast. Nitrogen fixation, efficient mineral uptake, and stress resistance all have an energy cost. The relation of benefits to costs in these processes must be optimized. This can only be based on an understanding of the chemical and physiological details of how the intact plants manage these interactions. The methodology for this understanding is available or nearly so, and thus it is simply a matter of applying more effort to solve the problems.

Mineral Uptake, Deficiencies, and Excesses

As a crop plant interacts with its soil environment, a number of mineral stresses can impair its functions. While these stresses are clearly recognized, they are poorly understood. However, the tools are available for in-depth examination of these problems. Single-gene mutants that vary in response to mineral stress are known and could facilitate investigations into the fundamental mechanisms underlying mineral-related problems and their potential solutions. Increased research in this area would provide the knowledge needed to improve crop production and efficiency in affected environments.

A mineral stress can be imposed by too little of an essential nutrient, too much of a toxic element, or the imbalanced distribution of a nutrient within a plant. The major essential nutrients, nitrogen, phosphorus, and potassium, must often be applied to agricultural soils at great expense in terms of both money and resources. Additionally, some soils are also deficient in micronutrients (zinc or boron, for example). The efficient utilization of a nutrient is a function of the uptake, distribution, and ultimate utilization of the molecule. Current understanding of these processes is sketchy at best. That there is genetic variability in nutrient uptake and utilization is recognized, but not well understood. While a number of instances of single-gene control of differential efficiency of nutrient utilization are documented (for example, iron utilization efficiency in corn and tomato), neither the basis for the differential efficiency nor the nature of the gene or genes that control this efficiency is known. Mycorrhizal associations are significant factors in mineral nutrition, but much remains to be learned about the nature and regulation of these associations. (See section on Organisms in the Rhizosphere for discussion of mycorrhizae.)

The benefits to be derived from a greater understanding of nutrient utilization are twofold, namely, the more efficient production of crops on soils currently supplemented with high nutrient inputs (particularly nitrogen) via a reduction in inputs, and the extension of cropping to marginal soils deficient in essential nutrients. These improvements might be gained either through the development of genotypes that more efficiently utilize low levels of nutrients or perhaps by improved management practices that might result from a better understanding of the nutritional aspects of the agricultural ecosystem.

Mineral toxicities occur worldwide. In semiarid regions, salinity is a serious limiting factor. In the tropics, aluminum is the most serious mineral toxicity. Salinity affects perhaps 30 percent of the world's arable soils, and aluminum toxicity affects approximately 40 percent. In the United States, salinity affects many soils in the Southwest, and aluminum toxicity occurs primarily in the Southeast.

Similar knowledge is needed with regard to both of these major types of soil toxicities. For the most part, the fundamental mechanisms by which these toxic influences affect plant processes are little understood. As with nutrient efficiency, genetic differences in tolerance to these toxicities are well known but poorly understood. A number of mechanisms for tolerance are postulated. These include, for example, rhizosphere modification, ion exclusion, chelation, and compartmentalization. However, extensive investigation is needed to identify key mechanisms for particular crop-ion combinations. The energy cost of tolerance mechanisms is another consideration that bears investigation.

Among the important mineral toxicities, salinity has been a major subject of increased research activity. Recent genetic and physiological studies have yielded valuable insights. Comparable investigations into aluminum toxicity (and associated stresses in tropical soils) will also be productive. However, much more extensive study is needed in both areas before the nature of these stresses can be clearly understood.

The applications of knowledge about mineral toxicities are similar to those for mineral efficiency. This knowledge could be used to develop new crop genotypes better able to tolerate these toxicities. Such genotypes would permit efficient crop production on affected soils now under cultivation by reducing high-cost inputs (e.g., water leaching for salinity, liming for aluminum toxicity). Additionally, such genotypes could also be used to bring marginal land into production, thus extending the range of arable land. Increased knowledge might also result in improved cultural practices that might be introduced in areas affected by these stresses.

There is evidence that some disorders in crop plants involve deficiencies of nutrients--such as calcium--in specific organs, even though the overall concentration of the particular nutrient in the plant may be adequate. Recent findings indicate that calcium nutrition interacts with a number of other important influences on a crop (e.g., a large number of physiological disorders damaging to many fruits, root crops, and leafy vegetables during storage are related to their calcium content, and calcium also influences tolerance to salinity and aluminum toxicity). These findings underscore the importance of further study in this area.

Biological Nitrogen Fixation

Research productivity in biological nitrogen fixation continues at a dynamic pace in chemistry, biochemistry, genetics, biology, and agronomy, with every indication that this pace will continue. Examples of recent research advances include sequencing of the nitrogen-fixation structural genes of several organisms, identification of regulating genes of nitrogen fixation, establishment of the similarity of the promoterlike sequence of algal nitrogen-fixation genes to plant chloroplasts, characterization of the molybdenum-iron cofactors of nitrogenase, identification of a nitrogen-fixing microsymbiont with nonlegume trees and the use of this microsymbiont Frankia to inoculate trees as the symbiotic bacterium Rhizobium is used to inoculate many legumes, increased yield of crops in specific geographic areas following inoculation with the bacterium Azospirillum, the enhancement of nutrient-ion uptake by Azospirillum suggesting opportunities to improve scavenging of limiting nutrients by crops, and the breeding of alfalfa-Rhizobium combinations with increased nitrogen fixation.

Nitrogen is a key limiting nutrient for crop productivity, with biological nitrogen fixation currently supplemented by 50 million metric tons of fertilizer nitrogen annually. Based on current technologies, it is estimated that the fertilizer nitrogen supplement will need to grow to 160 million metric tons by the year 2000. The fossil energy required to operate and the capital needed to construct fertilizer plants will be major expenses for agriculture. Applications from the expanding base of knowledge in biological nitrogen fixation may reduce the projected expansion in fertilizer nitrogen. Examples of highly probable short-term impacts are more effective *Rhizobium* inoculation technologies; legumes that are yield-responsive to fertilizers, thus increasing the harvest index for nitrogen; improved energy-efficient *Rhizobium* with genes to recapture some of the hydrogen that occurs as a by-product of nitrogen fixation; and *Azospirillum* inoculation of cereals to increase yields by presumably improved nutrient scavenging. Examples of less probable, midterm-impact technologies include *Rhizobium* that nodulates with insensitivity to the presence of high levels of fixed nitrogen and promiscuous *Rhizobium* to simplify inoculation. The more speculative, long-term impact technologies include achieving a major increase in photosynthate to fuel the symbiotic nitrogen-fixation system, finding catalysts that will produce synthetic nitrogen fertilizer with a zero direct energy input process, creating effective non-*Rhizobium* nitrogen-fixing associations, extending *Rhizobium*-legume-type symbiosis to nonlegumes, and transferring functioning nitrogen-fixation genes directly into crop plants. The latter is one of the most challenging objectives of genetic engineering. The magnitude of the need for fixed nitrogen justifies the search for multiple alternate technologies.

PLANT GENETIC ENGINEERING

Initially, the most important contributions of the application of genetic engineering techniques to crop improvement may come not from splicing and recombining specific traits into crop plants, but rather from the enormous potential these new tools offer for acquiring a better understanding of the basic biology of plants.

The engineering aspects of genetic engineering consist of the biotechnical innovations that permit scientists to isolate genes from cells in the form of DNA molecules and to splice them into the DNA of other cells. As genetic information in the form of DNA is common to all living systems, the kinds of techniques developed and demonstrated with the DNA of bacteria and viruses also are applicable to the DNA of plants and animals. In effect, genetic engineering provides a new technique for recombining genetic information in plants by permitting the recombination to be done at the molecular level. This new biotechnology of genetic engineering is a powerful adjunct to plant breeding, which hitherto has been restricted by the limitations associated with manipulating whole organisms.

To identify and isolate a gene in the form of a DNA sequence, to clone copies of this DNA that can be inserted into a living cell, and to provide for the expression of this foreign DNA involve an enormous scope of effort. Exciting advances, all successfully demonstrated, include the isolation of plant genes, their characterization by DNA sequencing, their duplication through cloning, and their insertion as replicating foreign DNA into other plant cells. It is probably only a matter of a short time before the expression of a foreign, recombinant, plant-DNA gene is definitely demonstrated.

In spite of the significance of the foregoing developments, the genetic engineering method for transferring genes will not easily replace genetic recombination that can also be accomplished through conventional plant breeding. What, then, is a realistic assessment of the value of genetic engineering to agricultural research and to the improvement of crops?

To begin with, molecular biology and genetic engineering techniques have already provided unexpected details of and new insights into the organization and function of genetic traits. The direct demonstration of relationships between regulatory and structural genes, the presence of repeated sequences of DNA that occur with high frequency throughout the genome, the existence of noncoding sequences (introns) that must be processed out of the RNA transcript, and the surprising plasticity of the genome that was earlier thought to be rigidly conservative have all helped and forced scientists to think in new ways about the factors that influence and control the reassortment and expression of hereditary traits. Much of this increased understanding could only be derived through the use and analysis of the technologies associated with genetic engineering.

In terms of the direct application of genetic engineering to crop improvement, it is not possible to predict what specific practical advantages genetic engineering techniques will have over conventional breeding techniques in the near future. The application of this new technology to plants is still very much in its infancy, and there are many uncertainties about how crop species will respond to manipulation using the new biotechnology. It seems reasonable, however, to expect that there are traits controlled by a single gene or small groups of genes that can be transferred successfully using genetic engineering. Such transfers, which would otherwise not be possible through conventional breeding, might involve the exchange of DNA between noninterbreeding populations of plants or perhaps even between plants and viruses or bacteria. The end result would be to enlarge the genetic diversity available for crop improvement. This appears most probable for single-gene traits that convey resistance to herbicides and to some diseases, or tolerance to adverse soil conditions such as mineral deficiencies or excesses. The genetic engineering transfer of simple disease-resistance traits, however, seems impractical considering the rapidity with which pathogens regain virulence against single-gene resistance.

A major uncertainty in genetic engineering applications to plant improvement programs lies in the fact that yield and many of the desirable attributes of cultivated plants are conditioned by multigene systems in which the genes interact additively and epistatically. Epistatic genes alter the phenotypic expression of other, nonallelic genes. At present, the functional control of multigene systems is not understood. The optimization of such complex multigene systems for their most effective functioning may be beyond the current techniques of genetic engineering. Knowledge of the function of these multigene systems, however, is likely to be greatly benefited by molecular studies of gene expression.

Some groups of coordinately regulated genes are known to be organized in multigene families such as those controlling the synthesis of the storage protein in some seeds. While it represents an enormous task, some understanding of the biochemical functions encoded by such genes seems a prerequisite to the ability to enhance the characteristics that are determined by these multigenic systems.

Novel genetic variability can be generated by means of several plant cell culture techniques. Mutant selection in plant cell cultures has already yielded a number of new genotypes of interest, particularly in the area of herbicide resistance. Such mutants may have utility not only in agricultural production, but also as highly selectable markers for other genetic and molecular studies. Somaclonal variation (the unselected genotypic variation sometimes observed in plants regenerated from cultured cells) has already provided valuable new germ plasm in the case of sugarcane, and promises to do the same in other crops. The fundamental basis for somaclonal variation is currently unknown, and further study of this phenomenon may well reveal much about the nature of the plant genome.

Both mutant selection and somaclonal variation are technically simpler processes than is gene transfer. The technology required is less complex, and the difficulties associated with introducing a gene into a foreign setting are avoided. For certain breeding objectives, these two processes may be the preferred means by which to obtain the desired variability.

The assessment of the potential contribution of new genetic technologies in relation to that of conventional breeding will vary from crop to crop. For several major agronomic crops, very sophisticated and powerful conventional breeding technologies exist. But some important crops are very difficult to manipulate by conventional genetic means simply because they are difficult or impossible to reproduce sexually (sugarcane, banana, bamboo), or because their life cycle involves a long delay in reproductive maturation, which is true of most woody perennial crops. In these species, simple improvements that might otherwise be unattainable, in practical terms, might be achieved through genetic engineering and plant cell culture techniques.

The successful application of genetic engineering will generate new and previously unavailable sources of desirable genetic variability that can be exploited through conventional plant-breeding techniques. It has the potential for providing genetic variability that otherwise might not be obtainable through induced mutations or exotic races of the same plant species. Thus, interactions should be encouraged between scientists working with conventional plant-breeding techniques and those engaged in genetic engineering research in order to facilitate the interchange of ideas and material. It should also be kept in mind that even with many crops whose yields have been markedly increased by plant breeding, additional yield improvements still appear to be possible through conventional plant breeding, and that innovative adjuncts to conventional plant-breeding methods are still being discovered--for example, the use of meiotic mutants in Solanum species (potato) that result in unreduced gametes, which facilitates crossing potato species with different ploidy levels.

And finally, the positive side of the unpredictability of research in various areas of genetic engineering needs to be emphasized. The unexpected is part of the very nature of any discovery. One hundred years ago no one could have predicted the configuration of American agriculture today because no one could have predicted the scientific developments that would help shape that configuration. New knowledge acquired through advances in genetic engineering, together with the factor of nonpredictability, could likewise lead to major new developments in modern agriculture in the years ahead as research pursues its sometimes anticipated, sometimes unforeseen course. The knowledge of the biology of plants that is to be gained by aggressively exploiting these new research tools is a valuable end in itself.

REPORT OF THE RESEARCH BRIEFING PANEL ON NEUROSCIENCE

Research Briefing Panel on Neuroscience

Floyd E. Bloom (Chairman), Arthur V. Davis Center for Behavioral Neurobiology, Salk Institute
Jack D. Barchas, Stanford University Medical Center
David H. Cohen, State University of New York, Stony Brook
John E. Dowling, Harvard University
Glen R. Elliott, McLean Hospital, Boston, Massachusetts
William Hodos, University of Maryland, College Park
Bruce S. McEwen, Rockefeller University
Charles Nicholson, New York University Medical Center
Donald J. Reis, Cornell University Medical School
Richard J. Wurtman, Massachusetts Institute of Technology

Staff

Fredric Solomon, Director, Division of Mental Health and Behavioral Medicine, Institute of Medicine
Allan R. Hoffman, Executive Director, Committee on Science, Engineering, and Public Policy

REPORT OF THE RESEARCH BRIEFING PANEL ON NEUROSCIENCE

For this briefing, neuroscience was defined by the Office of Science and Technology Policy as "that body of research directed toward understanding the molecular, cellular, and intercellular processes in the central nervous system (CNS) and the way in which those processes are integrated in CNS functional control systems, with emphasis on research relating CNS functions with behavior."

Neuroscience is a relatively new and rapidly growing investigative field. One index of this growth has been the almost linear increase in members of the Society for Neuroscience--from 250 in 1971 to nearly 8,000 today. Programs of education in the neurosciences at the undergraduate, graduate, and post-doctoral levels have increased by 200 to 300 percent over the past four years. This growth has provided a sizable pool of trained investigators ready to pursue the opportunities for discovery within neuroscience.

Neuroscientists share overlapping interests, concepts, hypotheses, and methods of analysis. The field transcends traditional disciplinary boundaries and now includes many workers who have chosen to move out of their original scientific niches and into studies for which an interdisciplinary approach bridges many communicative gaps and allows for more rapid technical and conceptual discoveries. Given the pace of discoveries in this field, with its cascading flow of information, neuroscience has been recognized as an exciting field of fundamental investigation. The ultimate goal of neuroscience investigation is to understand the human brain. The opportunity to investigate seriously the physical basis of certain forms of mental activity is, thus, beginning to break down many of the traditional barriers that have existed between the biological and behavioral sciences.

Neuroscience research has enormous implications for health. The brain itself is a major potential site of diseases, such as the major neurological and psychiatric disorders. Moreover, the brain can also be the origin of disease states expressed by the rest of the body by generating behaviors, such as smoking or obesity, that can produce or exacerbate lung cancer, diabetes, or heart disease. Disorders due to brain damage--e.g., sensory impairment, convulsive disorders, mental retardation, stroke, and dementia--now afflict more than 25 million people. Another 2.5 million people suffer from severe mental disorders such as schizophrenia or major depression. Furthermore, the United

States Public Health Service has estimated that at least 50 percent of the mortality from the 10 leading causes of death in the United States result from aspects of individual behavior. The health care burden of behaviorally related problems continue to accumulate, because many of involved diseases are chronic, rather than fatal. Conservative estimates of these costs run to tens of billions of dollars each year.

Because dysfunctions of the brain's normal regulatory controls are major sources of illness, support for neuroscience has spread across many of the federal health research agencies. Lead agencies in funding neuroscience research are the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alcohol, Drug Abuse, and Mental Health Administration (ADAMHA); other agencies providing important support for neuroscience include the National Institute on Aging, National Eye Institute, National Institute of Dental Research, National Aeronautic and Space Administration, National Science Foundation, and Department of Defense. Overall, the 1983 level of support for neuroscience research is estimated to be less than $500 million, a dollar amount that offers less buying power than existed in the federal budget for this field a decade ago. The current level of investment is small in relation to the potential of such research for reducing the burden of illness and the specific scientific opportunities that exist now for accelerating progress.

TECHNICAL ADVANCES IN THE NEUROSCIENCES

This report describes some major opportunities for solving important problems in neuroscience. We first review some of the more recent technical and conceptual advances that now make these problems amenable to productive research. Other newly emerging methods, on the verge of wide application, also hold great promise, but the present report is a state-of-the-art view of a field in which important ideas have consistently followed the development of new investigative technologies. These advances are described according to the four major domains of inquiry in neuroscience research: molecular, cellular, intercellular, and behavioral.

Molecular Research

Modern methods now permit the isolation and identification of specific molecules used by nerve cells to maintain their individual function as well as molecules that enable them to communicate with other neurons. Powerful new techniques have become available for investigating the molecular composition of the nerve cell membrane. Methods ranging from highly specific antibodies to physical probes of membrane fluidity provide information on molecules that are specific to certain classes of brain cells. Recombinant DNA techniques have recently been used to characterize the gene sequences responsible for neuron-specific products, especially certain of the neuropeptides that appear to transmit signals between nerve cells (neurotransmitters). The same methods may soon reveal the structure of their receptor molecules and of other components unique to nerve cells. The design of better molecular probes to provide a more complete chemical description of the nervous system is under way.

New analytical chemical techniques offer marked improvement in detecting, characterizing, and measuring substances within the nervous system. Such methods as mass spectrometry, nuclear magnetic resonance, high performance liquid chromatography, and radioimmunoassays can be used alone or in combination to detect substances at one-tenth to one-hundredth the concentrations needed with earlier methods.

Available evidence suggests that only a fraction of the neurotransmitters used by the brain are now known. These generally were discovered by seeking in the mammalian brain substances that were first encountered in other tissues, including the gastrointestinal system, the endocrine system, and neurons of invertebrates. It now is clear that an unexpectedly rich variety of chemicals are used for communicating within our brains--repeated time and again from strain to strain and from organ to organ. Neurotransmitters are now being discovered at a rapid pace due to the perfection of classical methods for isolating biologically active substances from tissue extracts and by the more recent application of the methods of modern molecular genetics.

The molecular mechanisms underlying the electrical activity of nerve cells also is being examined more thoroughly with the newer biophysical methods. Until recently, virtually all such information was derived from the use of microelectrode probes. Although much was learned from such studies, there were severe limitations in their ability to detect gradual electrical changes in the cell that generate neuronal action potentials. Furthermore those electrodes were especially difficult to use on small neurons, which are the most numerous cell class in the brain. Within the past few years, a new recording technique has been been developed as a result of advances in solid state electronics. It promises to provide information on membrane properties thought virtually unobtainable a few years ago. This technique, commonly referred to as "patch clamping," enables an investigator to record from the tiniest of neurons and detect the opening and closing of a single ion channel as a small patch of membrane responds to neurotransmitters. The method also enables the investigator to manipulate the internal environment of a single neuron. As more aspects of this technique are perfected, the data already available indicate the beginning of a new magnitude of discovery.

Cellular and Intercellular Research

Much of the current excitement in neuroscience comes from an increased ability to understand the biology of the cells of the brain and the molecular processes by which they communicate and regulate each other. New methods for tracing circuits in the brain and for determining the molecular mechanisms underlying the transfer of information between nerve cells have greatly enriched our understanding of how the brain receives and processes information. However, the task of describing the connections of the brain and the actions of these connections is far from complete. To be fully appreciated, the newly identified circuits also must be considered in terms of the transmitter substances by which they communicate and the molecular mechanisms that mediate these signals. These aspects of intercellular communication are now capable of being studied together.

Expanded understanding of a biochemically specifiable cellular neuroanatomy has not only enlarged our views of how neurons communicate but has also permitted the development of new approaches to assessing and treating diseases of the brain and behavior. With every new transmitter isolated and identified, research provides a means to make analogs that can simulate or antagonize their actions, as well as methods to measure their relevance to unexplained diseases of the nervous system. Such studies also are far from complete.

Much useful information on the cellular properties of the nervous system is being gained through the use of cell cultures and other _in vitro_ preparations. These are uniquely suited for studying such important issues as the definition of chemical and environmental requirements for neuronal maintenance and neuronal growth.

Genetic mutants are being used increasingly in analyzing the functional role of specific neurochemicals or structures and in modeling certain disease states. Typically, mutants are selected on the basis of a behavioral or neurological deficit and then studied for their underlying chemical or structural defects. For example, mice showing major deficits in postural orientation and locomotion have been shown to lack the normal cellular organization of a major brain structure, the cerebellum. Mice lacking receptors for the male hormone testosterone are found to be deficient in the ability to display certain male-specific behavior and to lack certain male-characteristic neural structures, notably in the spinal cord nuclei that normally innervate the penis. The process of genetic hybridization is permitting the development of animal strains with specific neurological or behavioral properties, ranging from resistance to alcohol intoxication to the ability to acquire new reflex skills. Such strains increase the chances of determining the cellular mechanisms underlying genetically expressed traits.

Behavioral Process Research

In recent years the available repertoire of techniques for recording the activity of the brain during behavior has expanded greatly. Methods now range from the study of the activity of single nerve cells in animals to the ability to describe regional metabolic activity in human beings. Applications of these powerful methods are opening a new understanding of the neural basis of certain behavior. It now is possible to correlate the activity of specific neurons with such well-controlled behavior as learned and unlearned movements of the hand and the eyes, and thus to gain a high-resolution view of how specified arrays of neurons process information in the initiation and control of patterns of muscular contraction. In experiments for which an animal is taught to make or alter some behavioral response, studies now reveal the critical value in this process of systems by which the brain selectively directs its attention to the events of the internal or external world. This type of study begins to explore the changes in neuronal activity that reflect learning and the storage of information.

New research tools include such techniques as recording the brain's electrical activity without having to insert electrodes directly into the brain (so-called "noninvasive recording"). The development of imaging techniques, such as Position Emission Tomography (PET Scan) and Nuclear Magnetic Resonance scannings and of other direct functional measurements, such as event-related electrical potentials, have opened new opportunities for approaching questions about the neural localization and basis of higher cognitive function and perception in human beings. As normal patterns of brain activity are characterized, the potential of these techniques for detecting and identifying abnormal function increases dramatically. Remarkable advances in this context are at their inception. The capacity to span investigations from single-cell correlates of behavior to correlations between populations of neurons and behavior offers promise for an understanding of how arrays of neurons are organized to produce the ultimate output of the nervous system--behavior--and for exploring how such arrays of cells can be modified by experience and perturbed by certain environmental influences and disease.

New concepts and techniques in the analysis of behavior have made essential contributions to efforts to study basic neurochemical and neurophysiological mechanisms. Increased precision in recording of behavior has vastly improved studies of its neural and hormonal controls. Technical advances also have contributed to the better analysis of behavior. For example, telemetry with battery-powered, portable transmitters has made possible the study of freely moving animals in the wild and their patterns of social behavior. Similarly, video recording, which enables later multivariable scoring by several different observers, has led to sophisticated analyses of behavior ranging from aggression to mother-infant interactions.

Computers in Neuroscience

The advent of powerful, inexpensive microprocessors and increased access to large computers has opened new avenues of research at all levels of inquiry in neuroscience. For example, anatomical data gathered with a variety of microscopic techniques, including precise tracking of specific molecules in selected cells, can be assembled into data bases in a computer and then reconstructed into three-dimensional representations and displayed as color graphics. In other applications, large populations of suitably labeled cells can be identified and quantified with computer techniques, thus enabling comparisons among normal, experimental, and disease-modified brain regions. Computers also make feasible the collection and analysis of simultaneous recordings of electrical activity of many neurons. On-line microprocessors are particularly useful for condensing, displaying, and storing such data. This immediate analysis of data can guide the experimenter in the choice of stimulus-response tests during the experiment.

Recent advances in defining aspects of behavior, together with video recording, permit a quantitative approach to behavior. In this context, the computer can be used to administer and modify behavioral models and to aid in codifying and collating data. In the latter context, video images can be analyzed frame by frame using interactive computer-driven graphic techniques, thus permitting very large assemblies of data to be handled.

The rapidly expanding availability of small computers and terminals, together with inexpensive data communication capabilities, makes feasible the establishment of networks of frequently updated data bases in many areas of neuroscience that groups of investigators can share as common repositories of information.

TARGETS OF OPPORTUNITY IN NEUROSCIENCE

Interest in neuroscience arises from efforts of many scientists to explore certain basic questions about the brain, its growth and function, and disorders that alter its ability to function properly. The opportunities selected for attention in this report relate to specific aspects of the same questions that have aroused the curiosity of scientists of the brain for decades. What now opens these questions to serious investments of experimental effort is the emergence of adequate tools, an accumulated background of fundamental information, and the trained manpower to make these efforts fruitful. Eventual success of the effort depends on the maintenance of research programs in the fundamental subdisciplines of neuroscience from which the present exploratory opportunities will be launched. Five major opportunities, selected because of these recent developments, offer extraordinary promise for promoting an understanding of the structure and function of the behaving brain, and its relevance to major health problems.

How Does the Brain Grow and Maintain Itself?

Contemporary neuroscience is now in a position to consider how the brain assembles itself from its primitive neuronal elements, how the neurons differentiate with specific biophysical and biochemical properties, how they seek out and connect with other neurons to form the adult brain, and how the neurons maintain themselves and their connections. Current efforts seek to establish whether neurons keep an ability to form new connections in adult life, either in response to environmental signals or to injury, and whether these properties are lost as neurons age. It now seems possible to determine how the orderly process of growth and differentiation is regulated by the fetal environment, by maternal factors, and by the mother's health and environment. This research is moving to the point of determining how external factors, interacting with gene expression, result in abnormal patterns of development and how such abnormalities result in the congenital brain defects affecting individuals with mental retardation or eventual mental disease. The field is in an explosive growth phase, not only for neuroscience but for biology in general. Three representative topics may be singled out: growth, differentiation, and maintenance of neurons.

It was once believed that neurons formed, made connections to each other, and developed into the adult brain in an orderly and genetically preprogrammed way. There is now evidence that these processes depend on or can be modified by chemical agents secreted by other cells and sensed by the neurons. Some factors are essential for neuronal survival when connections are first being made. Other factors may direct gene sequences needed to produce specific neurotransmitters. Still others may stimulate and direct the growth of neuronal processes to specific target cells. At present only a few factors have been identified, purified, and synthesized. Their further identification is hampered by their existence only in very small amounts. The techniques of recombinant DNA technology should further promote the identification and synthesis of such compounds and permit analysis of how they may influence the normal growth and development of nerve cells.

Neural Regeneration and Transplantation

The fact that mature nerve cells do not undergo cell division makes the brain particularly vulnerable to injury. Repair in nerves outside the brain can be achieved by regrowth of the neuronal processes. For example, severed nerves in the hand can undergo full regeneration. It has recently been recognized that cells of the brain also have the capacity to regenerate connections; however, some unknown factors apparently interfere with such growth. It has been proposed that such processes of regrowth may underlie the partial recovery of function that follows brain injury, stroke, or spinal cord damage. Present knowledge suggests that other specific tissue factors may promote growth of neurons. Additionally, in some experimental animals, certain neurons lost by degeneration or by congenital absence can be replaced by transplantation of comparable neurons from other normal animals. A preliminary attempt has already been made in Sweden to apply this transplantation approach to treat a patient with Parkinson's disease. These opportunities rest upon the development of anatomical techniques for identification of chemically specified nerve cells, improved techniques for studying regulatory communicative processes in tissue slices, and the increased accessibility of the techniques of molecular genetics for identification of small regulation molecules. We now have the opportunity to acquire further information about how such transplants can replace functions of absent cells in the brain, and possibly provide eventual applications to treating human disease.

Aging

Aging relates to the brain in two ways. First, some populations of nerve cells seem especially susceptible to dying with age, leading to the clinical features of the aging brain, dementia, and loss of mobility. The bases for these cellular losses are unknown. A second, more speculative relationship is that loss of a small population of brain cells, possibly endocrine-related neurons of the hypothalamus, may be the biological trigger that results in the intrinsic aging of tissue throughout the body. Perhaps one of the most promising areas in neuroscience relates to dementia, a problem destined to increase in social importance as the proportion of our aged population increases. Recent studies indicate that a specific group of cellular connections between the forebrain and the cerebral cortex appears to be lost in individuals with senile dementia. The transmitter for this circuit is believed to be acetycholine. When drugs that interfere with acetylcholine transmission are given to normal subjects, many of the memory defects characteristic of the demented individual are produced. These findings are of great importance, for they may relate a specific cell population with a known transmitter to a disease of important economic consequences. The neurochemical identity of these cells raises the possibility of identifying subjects at risk by analysis of body fluids and offers the prospect of designing drugs that could replace actions of the lost neurons. Yet to be discovered are the molecular properties of these cells that make them particularly vulnerable to degeneration, the role of genetics in predicting vulnerability, and the possible role of environmental agents in the degenerative process.

How Does the Brain Acquire, Store, and Use Information?

An important aspect of the brain's role as the organ of behavior is its ability to acquire information about the external and internal environments, to store and evaluate this information, and to take specific action based on it. Recent advances in the field of sensory motor integration enhance our understanding of how the brain regulates the flow of information it receives. Advances have been made possible by studies of events in the brain cells of awake animals engaged in monitoring and vigilance tests; of the precision of eye movement in animals and humans; and of attention, sensory mechanisms, and perception in animals and humans.

Other studies appear to be pointing the way toward solving the mysteries of how and where information is stored in the brain. Progress is being made on several fronts. At the cellular level, membrane specialists and neurochemists have begun to unravel the nature of the modifications that occur with time after a stimulus is received. There have been studies of memory and its neural correlates in animals ranging from invertebrates to nonprimate mammals, whose nervous systems are far less complex than those of human beings. Observed changes suggest that experience modifies neuronal communication through specific metabolic molecular events that control the physical properties of the nerve membrane and the secretion of neurotransmitters. Using the new culture techniques, it has been possible to begin studying the storage capabilities of neuronal networks contained within functional, isolated slices of mammalian brain. Other studies have investigated learning, memory, storage, and retrieval in animals, in normal children and young adults, and in people who have various types of memory disorders, such as the normal aged and those with Alzheimer's disease.

Two other promising subdisciplines deal with the ways in which the brain uses information. One of these is neuroethology, which studies the neuronal basis of behavior related to an animal's or person's interaction with the environment and with other organisms, such as courtship, parental care, establishment and defense of territories, aggression, and social communication. Many of these behaviors are genetically preprogrammed, but they also can be affected by specific experiences. The other is the field of cognitive neuroscience, which investigates a wide range of intellectual processes in normal and neuropathological human populations. This approach has a range from the study of event-related electrical activity in the brain in awake humans to studies of language, decision making, and problem solving.

How Does the Brain Monitor and Regulate the Body?

In addition to generating mental activity, thoughts, and feelings, the brain also regulates the levels of key chemicals in the bloodstream and thereby influences the health of every cell in the body, including its own neurons. The brain receives chemical information from specialized sensor cells whose outputs vary with the concentration of such compounds as oxygen, glucose, or certain hormones, or with such physical variables as blood pressure or body temperature. Special groups of neurons then "decide" whether this component is appropriate and, if not, activate nerve cells capable of restoring normal homeostasis. Such accommodation occurs in any of three output channels: 1) neuroendocrine transducers, which release hormones into the blood stream or actively inhibit their release; 2) autonomic neurons, which control such events as the beating of the heart, the resistance of the blood vessels, the release of energy-rich compounds from the liver and muscles, and the secretion of some peripheral hormones; and 3) behavior, which might call for drinking water to restore blood volume, putting on more clothing to raise body temperature, or eating food to raise blood glucose. Disturbances in these sensor-integrator-effector loops may lead to a number of common disorders, including hypertension, sudden cardiac death, obesity, and infertility. An understanding of which brain neurons are involved--of their locations, their neurotransmitters, connections, and their responsiveness to circulating hormone levels--is likely to facilitate the development of effective treatments for these diseases.

The brain is selectively permeable to many circulating compounds other than hormones and key metabolic constituents. Recent evidence suggests that this permeability can be altered, with adverse effects, under circumstances such as uncontrolled diabetes or severe kidney or liver disease. An understanding of the dynamic relationships between blood and brain levels of nutrients and metabolites should provide insights into such diseases and into brain function in general and suggests ways to take advantage of these relationships for the treatment of the brain dysfunction.

The brain also receives a great deal of information from the lungs, heart, blood vessels, and gastrointestinal tract through nerves. Although the existence of these nerves has been known for a long time, much remains to be learned about the central interconnections and the neurotransmitters involved in their communication. It is now clear that disturbances in their functional control by the brain can lead to major medical problems, including asthma, sudden infant death, hypertension, sudden cardiac death, and possibly peptic ulcer.

Another finding in recent years is that the body's immune system is not as independent of the brain as previously believed. Psychosocial events such as loss of a spouse or other severe stresses can result in a temporary depression of immune responses to infection or foreign agents in the body. There are suggestions that these effects are mediated through the brain, and it should now be possible to identify the hormones, neurotransmitters, and neural systems involved in modulating the immune system.

How Does the Brain Express Rhythms, Drives, and Emotions?

Certain behavioral functions of the brain, such as sleep, arousal, aggression, sexual drives, eating, drinking, and even drug-seeking behavior, are now subject to serious investigation. Progress in understanding some of these already is substantial. For example, at least in animals, neural pathways related to sexual behavior, aggression, and sleep have been mapped; some of the relevant neurotransmitters and sites of hormone action also have been identified. Recent discoveries of a number of previously unsuspected neurotransmitters raise questions about their relevance to these behaviors and states. The discovery of brain substances (endorphins) that have properties like those of opiate drugs has led to a broad new line of inquiry into addictive states.

Hormones secreted by the endocrine glands are regulated by the brain through signals from the pituitary. In turn the endocrine hormones feed back to the brain evidence of their activity, which permits the brain to keep the system within normal fixed limits of activity. However, these same hormones also can influence mood and motivate behaviors; such effects, still incompletely explained, may influence the response of the brain to drugs used to treat mental and neurological disorders. Sex-related differences in behavior and personality are now being traced to the hormonal actions to which the brain is sensitive early in development.

Even such everyday activities as eating a meal are now being subjected to more precise research by breaking down the behavior into its components. The sensation of hunger, the decision to eat based on sensory or social inputs, the choice of what to eat, and the decision of when to stop eating each may involve particular sets of brain neurons. Any of those neuronal sets could be responsible for appetite disorders, such as severe overeating or undereating.

Biological rhythms are another promising area of basic research that may be fundamental to understanding and classifying many types of inborn behavior. For example, the daily cycle of sleep and wakefulness normally is synchronized to the periodic external environment by an internal biological clock. This physiological timing system shares the properties of similar circadian rhythms in a wide array of life forms, from single-cell algae to the entire plant and animal kingdom. In recent years, there have been major advances in understanding the physiological and behavioral processes that affect the timing of the sleep-wake cycle and other biological rhythms. Such rhythms would be of interest in any case, but artificial lighting and rapid forms of travel have produced unprecedented changes in the temporal environment. The short- and long-term consequences of such shifts remain to be assessed adequately. The most important periodic environmental stimulus for entraining daily rhythms in nearly all species appears to be the alternation of environmental light and dark. A specific area of the brain has recently been identified as one pacemaker responsible for the generation of many of these rhythms.

An appealing feature of research on discrete, inborn brain states or drives is its potential application to more complex brain functions, such as memory and voluntary behavior. In relatively simple invertebrate animals, it now is possible to trace the entire neural circuit for such a behavior as orienting to a source of food. In theory, at least, investigators should be able to monitor the electrical activities of the cells in that circuit, to identify and study the neurotransmitters involved, and to study the effects on the system of specific environmental influences. Information gained from such studies should go a long way toward providing investigators with the tools needed to examine in detail more complex behaviors in higher animals, including human beings.

What Specific Aspects of Brain Function Are Altered in Disorders?

Rapid strides have been made in improving the diagnosis and treatment of some severe mental disorders, neurological conditions, and addictive states. The improving armamentarium of basic science has made it increasingly possible to examine these problems by means of human and animal models.

Biochemical Studies in Human Beings

Knowledge about recently discovered neurotransmitters will undoubtedly influence future advances in the understanding of psychiatric, neurological, and addictive states. Although further basic information is essential, the two dozen substances already regarded as important in communication between nerve cells make possible highly relevant studies of the role of specific neurotransmitters in various disease states. For example, investigators can assess the activity of various neurotransmitter systems by measuring substances in body fluids, which offers unprecedented opportunities for research into such disorders as depression and schizophrenia. Also, brain tissue taken after death is providing invaluable information about concentration and localization of neurotransmitters and related constituents, particularly in degenerative disorders of the CNS, such as multiple sclerosis.

Studies of the neurochemistry associated with neurological disorders and addictive states require not only substantial neurological and analytical chemical skills but also clinical research centers in which patients with various illnesses can be studied in controlled settings and with careful behavioral monitoring. Such facilities are rare, but their presence is essential.

Better techniques for studying brain functions also have clear implications for understanding other severe brain disorders. For example, Huntington's disease is an hereditary disorder characterized by involuntary movements and a deterioration of cognitive function. In most individuals with Huntington's disease, symptoms do not appear until well into their reproductive years, so that they must decide whether or not to have children without knowing if they will be transmitting this devastating disorder. Strides have been made in characterizing the anatomical and neurochemical changes that occur with this disease. Such information is hoped to help in developing early ways of detecting people who have Huntington's and in devising more effective treatments for its crippling effects. A similar strategy may be generalizable to other inheritable degenerative disorders of the nervous system that have more complex genetic and environmental contributions.

The new and still primitive noninvasive techniques for studying human brain function could lead to recognition of new categories of illness. Such methods are greatly expanding the potential for precisely localizing sites of brain damage and correlating such damage with changes in body function and in behavior. Available evidence already suggests that there are several forms of epilepsy, depression, and schizophrenia. In some cases clinically indistinguishable subtypes of disorders may prove to have quite different patterns of changes in the brain. There are suggestions that damaged areas in the temporal lobes of the cortex can present as abnormal behavior, such as episodic rage or psychotic thinking, rather than as classical epileptic seizures. Identification of such subtypes will be critical in designing more effective treatments, in developing strategies for prevention, and in formulating and testing hypotheses about the underlying causes of disease.

Other Information Needed on Human Problems

Additional fundamental information is required about behavior and about more traditional biology. Improved rating skills that for such specific components of mental activity as memory, speech, mathematics, emotionality, stress, pathological fears, or thought processes are important examples. An ability to rate changes through time, including quite subtle changes, will enhance studies of the natural history of disorders in these abilities and of their treatment.

Behavioral studies that tie into basic biological work, such as investigation of sociophysiology--the ways in which basic social processes alter physiology--may provide information that will result in better controls and more careful clinical studies, as well as being directly germane to studies of the disorders themselves. The need for animal models of mental, neurological, and addictive disorders continues to be pressing. Although human studies are essential, there are no substitutes for animal studies in many types of investigations. Thus, the recent development of an animal model for the sleep disorder narcolepsy provides a new opportunity to study a vexing and difficult process.

EMERGING TARGETS OF OPPORTUNITY

Aside from the major opportunities for advancing knowledge within neuroscience, there are certain areas of promise for which the base of fundamental information and available methods is not yet sufficiently advanced to predict definitive results over the next few years. In almost all cases, these emerging areas of opportunity need a longer term commitment because of the complexity of the problem area. In this category, we may mention the following.

1. What research strategies will clarify the currently unsolved neurological and mental disorders? Most modern concepts of CNS disease have rested on the hypothesis that some part of the brain's "hardware" has become pathologic--either a loss or perhaps a hyperactivity of a specific class of neuron. Recent advances in the ability to separate and purify the receptors of neurotransmitters and to follow their changes in the presence of chronic drug treatments suggests that there may be whole classes of "software" diseases which derive from abnormal patterns of receptor regulation. Also, greater insight into the nature of these receptors may well lead to the development of drugs that have been molecularly tailored to correct specific defects in cellular communication.

2. How does the brain deal with environmental toxins and drugs? Much has been learned about the ways in which brain cells respond to demands on their activity, but less is known about the ways in which toxic materials of various sorts produce their effects on the body, effects manifested in part or in whole through the brain. Especially critical are the effects produced by such substances as narcotics, alcohol, excess food leading to obesity, and cigarettes. These potential voluntary toxins account for a very large health burden, particularly costly because of the chronic nature of their abuse and the long lags between the adoption of the voluntary patterns of abuse and the appearance of adverse consequences to health. This includes possible effects on the developing fetal nervous system that may not be manifest by neurological or behavioral abnormalities until well into postnatal life.

3. How does the brain deal with infectious agents? Recent research has demonstrated that many degenerative disorders of the nervous system can be simulated in experimental animals by inducing immune responses against one or another neural component. These studies suggest that such responses could be triggered by the changes accompanying infection. Severe, lethal degenerative disorders of the brain are known to result from virus infections which occurred long ago, in which the virus has lain dormant in the nerve cells awaiting proper conditions for replication and tissue destruction. Some of the sites by which viruses such as the rabies virus can invade nerve cells may depend on the specific arrays of neurotransmitter receptors that a susceptible cell carries. Some recent data suggest a viral basis for some forms of age-related dementia and for certain rare types of schizophrenia. Other diseases for which virus-related clues have been raised, for example, multiple sclerosis, are not yet to this endpoint. Nevertheless, determination of the way in which the brain typically is protected from but can succumb to infectious agents is an area of immense importance for additional future investigation in which advances at the fundamental levels already in progress will be likely to pave the way.

SUMMARY

Neuroscience is a relatively new and rapidly growing investigative field that transcends the traditional boundaries of fundamental scientific disciplines and encompasses research efforts in all clinical disciplines relating to diseases of the brain and behavior. This briefing report has described some of the basic technical and conceptual advances that have fueled the recent explosive advances in neuroscience research. Given the availability of a critical pool of skilled investigators--steadily supplemented by new investigators in training and by the influx of experienced investigators entering the neuroscience field from other disciplines--important opportunities now exist for significant enhancement of research efforts. The Panel has selected five broad targets of opportunity that are likely to experience advances in the near-term future. These targets of opportunity are crystallized by the following five general programmatic questions:

1. How does the brain grow and maintain itself?
2. How does the brain acquire, store, and use information?
3. How does the brain monitor and regulate the body?
4. How does the brain express rhythms, drives, and emotions?
5. What specific aspects of brain function are altered in disorders?

The Panel believes that, with additional resources, the techniques and skills now available could be put to immediate and productive use on these broad concerns. The product of such increased efforts not only would yield important intellectual advances in man's ability to understand the function of the human brain but also would offer direct insights into the important unsolved neurological and behavioral disorders, which by their early onset and chronic course, represent an extremely costly burden of illness to the nation.

The skills and tools we have highlighted are far from being universally shared. The Panel has not attempted to address the practical realities of making available the recently developed, often expensive, investigative tools described, nor have we considered how scientists not yet engaged in neuroscience may be able to acquire the necessary research skills to enter the field. Although scientists in the field have expressed no consensus as to which areas may experience the next major advance, the five targets of opportunity to which this report calls attention represent areas of research to which the creative curiosity and individual efforts of many neuroscientists can be applied.

This report has emphasized areas that lie at the frontiers of neuroscience. It must not be forgotten that our present concepts of the structure and function of the brain have been built upon decades of painstaking, detailed research, during which time United States neuroscientists have been the leaders. Continuation of their fundamental efforts will be necessary if we are to avail ourselves of the opportunities that lie ahead.

REPORT OF THE RESEARCH BRIEFING PANEL ON HUMAN HEALTH EFFECTS OF HAZARDOUS CHEMICAL EXPOSURES

Research Briefing Panel on Human Health Effects of Hazardous Chemical Exposures

Norton Nelson (Chairman), New York University Medical Center
Bruce N. Ames, University of California, Berkeley
Earl Benditt, University of Washington
Theodore Cairns, Greenville, Delaware
Ronald Estabrook, University of Texas Medical School (Southwestern)
Bernard Goldstein, University of Medicine and Dentistry of New Jersey
Joshua Lederberg, Rockefeller University
Brian MacMahon, Harvard University School of Public Health
Richard Merrill, University of Virginia School of Law
Vaun A. Newill, Exxon Corporation
Gerald Wogan, Massachusetts Institute of Technology

Staff

Robert G. Tardiff, Executive Director, Board on Toxicology and Environmental Health Hazards, National Research Council
Warren Muir, Consultant
Allan R. Hoffman, Executive Director, Committee on Science, Engineering and Public Policy

REPORT OF THE RESEARCH BRIEFING PANEL ON HUMAN HEALTH EFFECTS OF HAZARDOUS CHEMICAL EXPOSURES

INTRODUCTION

This report is limited to research on the human health effects of hazardous exposures to chemicals, both man-made and natural. Although many physical factors, including ionizing and nonionizing radiation, are of public health importance, they are not included here. Nor did the Panel consider effects on the environment.

It should be recognized that personal habits, including cigarette smoking, unquestionably have major health impacts; and other factors, such as dietary constituents, may also be important. The initiatives proposed here should contribute a better assessment of the nature, variety, and extent of chemical threats, whatever their source, to the human population.

The extent of exposure is an indispensable component in any evaluation of the effects of hazardous chemicals. Research on means for estimating exposure by environmental monitoring, although very important, is beyond the purview of this Panel's study. One major proposal, however, deals with the estimation of individual exposure by the use of biochemical markers.

Many decisions made by the public and private sectors in relation to the control of risk of chemical exposures have had important economic consequences. Federal agencies and industry have often been compelled to make such decisions in the face of large scientific uncertainties. No amount of research can eliminate all uncertainties associated with assessing the risks of exposure to potentially hazardous chemicals or eliminate the controversial judgments inherent in any decision about control of chemical exposures. Yet there are opportunities to reduce scientific uncertainties and to increase the public's confidence that health is protected and that the economic consequences of imposed chemical control are justified. Recent dramatic advances, especially in molecular biology, have opened up the possibility of making substantial progress in reducing the scientific uncertainty in such decisions. Some of the best opportunities are of a fundamental nature that are likely to greatly enhance the ability to directly determine individual exposure, preclinical signs of disease, and chemically related illnesses in humans. Other promising efforts offer the prospect of greatly improved interpretation of laboratory data on possible chemical effects.

Of the many possible areas of potentially important and fruitful research, this Panel is recommending six that offer particular promise and are uniquely timely. Other important areas examined by the Panel included neurobehavioral effects, immunotoxicity, and organ-system studies, but they were eliminated from the Panel's primary recommendations at this time. These six areas were chosen based upon the Panel's belief that they offer the greatest prospect for making advances in our understanding of scientific issues that are likely to be pivotal in chemical risk decision making. Some of the initiatives relate to the improvement of existing assays; however, until such improvements are effected, existing methods must continue to be used. Some of the initiatives indicate that substantial progress can be made with research on several endpoints using human cells and tissues; however, the Panel recognized that several technical and legal considerations need to be resolved before such tools can be fully implemented and that for many other endpoints--beyond those covered in this report--substantial technical difficulties are anticipated before viable assays become generally available.

The six topical areas, for which priorities were not established, are:

I. Research to improve the basis for dose and interspecies extrapolations to humans.

II. Research into prediction of the effects of multiple chemical exposures.

III. Improvement in approaches to the determination of reproductive and developmental effects of chemicals.

IV. Development of cellular and molecular markers of exposure to chemicals and markers of preclinical effects.

V. Development and evaluation of means to distinguish among chemical cancer-causing agents on the basis of modes of action.

VI. Expanded use of existing federal data collection activities for assessing human health effects.

I. RESEARCH TO IMPROVE THE BASIS FOR DOSE AND INTERSPECIES EXTRAPOLATIONS TO HUMANS

NEED: Control decisions for human health are based largely on extrapolation from tests on laboratory animals. The qualitative and quantitative biologic differences between test animals and humans vary substantially. Therefore, improvement in making such extrapolation is needed.

RECOMMENDATION FOR PROMISING RESEARCH: Comparative studies--e.g., on enzyme patterns and repair mechanisms--that use, in the case of humans, surgical and autopsy tissues.

ANTICIPATED IMPACT OF RESEARCH: Research to establish the qualitative and quantitative differences over a wide range of doses between test animals and humans can improve toxicologic predictions.

RATIONALE: The information available to make judgments about the possible effects of exposures to chemical substances often consists only of results of laboratory tests on species other than humans. Judgments made on the basis of such information are inevitably weak, because of a lack of understanding of the factors determining interspecies similarities and differences in responding to chemical agents.

Interspecies variations in responses to toxic chemicals can arise from differences in many characteristics, such as genetic inheritance; enzymes that activate and inactivate toxic materials; intracellular pathways of toxicity; membrane biochemistry and receptors; absorption, distribution, storage, and excretion; and physiologic differences in particular organs. Through the study of these differences between humans and test species (including the use of organs and cells in culture), it should be possible to improve greatly our ability to interpret laboratory data in terms of potential risks to humans.

Exacerbating the problem of dosage extrapolation and interspecies comparison is the increased heterogeneity of the genetic mix as one proceeds from purebred, genetically homogeneous strains of mice to the human population. Genetic composition and expression are dominant factors in determining the cellular response to chemical insult. Knowledge of regulation of expression is still rudimentary, but some methods to evaluate differences in "genetic potential" are now available. Such methods must be used, if we are to recognize and evaluate differences in susceptibility arising out of human genetic diversity.

Improvement in the scientific base for dosage and interspecies extrapolation underlies improvement in the basis for judgments about chemical-related health risks to the human population. Such improvement is now possible through exploitation of increased understanding of a number of areas, including gene expression, DNA repair mechanisms, enzyme induction and specificity, developments in cell-culture techniques, and advances in cell biology.

A number of specific subjects of particular promise that should be pursued are noted below:

1. Differences in bioactivating and protective enzyme systems and of repair systems should be studied to increase understanding of differences in pathologic response between humans and test species.
2. Human and laboratory animal genetic homogeneity should be explored to gain the knowledge needed to recognize and evaluate differences of susceptibility arising out of human genetic diversity and to enhance extrapolation of laboratory results to susceptible members of the genetically diverse human population.
3. The role of physiological characteristics, such as lung structure and function, should be pursued in concert with molecular and biochemical explorations.

Thus, it is recommended that a focused program of research be undertaken to improve the basis for dose and interspecies extrapolations of chemical risks to humans.

II. RESEARCH INTO PREDICTION OF THE EFFECTS OF MULTIPLE CHEMICAL EXPOSURES

NEED: Most chemical hazards involve concurrent or sequential exposure to multiple chemicals, e.g., product mixtures, toxic wastes, cigarette smoke, and pollutants in air, water, and the workplace. The scientific basis for evaluating health effects of such mixed exposures at often widely differing dose levels is extremely primitive.

RECOMMENDATION FOR PROMISING RESEARCH: Study of mechanisms by which chemicals alter biologic responses to yield addition, inhibition, or synergism of effects.

ANTICIPATED IMPACT OF RESEARCH: Such research can substantially improve ability to assess the health effects of environmental chemical exposures.

RATIONALE: Although study and control of environmental chemicals have necessarily focused on individual compounds, it is obvious that humans are exposed to a multiplicity of agents in their daily activities. Most chemical products in commerce and nearly all environmental pollutants--whether in the workplace, from hazardous-waste sites, water-effluent streams, or formulated products--are complex mixtures. To the extent that the effects of chemical mixtures differ from those of their separate components, there is an important need to develop approaches to the prediction of the effects of the mixtures, so that the basis for making control decisions about exposures to them can be improved. The needed studies can be conveniently divided into two types: 1) those focusing primarily on interactions among several potentially hazardous environmental chemicals, and 2) those evaluating interactions among environmental chemicals and other agents, such as foods and drugs.

Toxic chemicals can produce independent, additive, synergistic, or protective effects. In some situations, not only the extent but also the type of toxic effect of several chemicals together differs from the sum of each of the chemicals acting alone. Combined exposures can result in effects in different organ systems or in different toxic effects. Our understanding of the mechanism of action of individual agents has in a few instances begun to allow the prediction of effects of multiple agents, but much more information is needed.

Considerable data are now available on the effects of exposures of laboratory animals to specific chemical agents with regard to their abilities to activate or deactivate other chemicals. Studies in this area should be extended to the human as an aid to elucidating the underlying sources of the varied sensitivities of persons exposed to toxic chemicals.

Future studies should emphasize not only the biotransformations that lead to interactive effects, but also other processes, such as the absorption and distribution of chemicals. For example, with respect to the effect of one chemical on the absorption of another, competition for receptor sites in cellular membranes may be important in hormonal mediation. Such processes should be studied to determine the circumstances in which competition may be important.

An overall emphasis in future studies, wherever possible, should be on developing an understanding of general classes of compounds as an aid in predicting the nature of interactive effects. Such studies should complement the research on dosage and interspecies extrapolation described in Section I.

The results of all these studies would be valuable in guiding decisions about the many complex mixtures in commerce and in the environment. They would be extremely valuable in establishing priorities for the cleanup of multichemical hazardous-waste dumps.

Thus, it is recommended that there be a substantial increase in research into ways of predicting the effects of multiple chemical exposures at various dose levels.

III. IMPROVEMENT IN APPROACHES TO THE DETERMINATION OF REPRODUCTIVE AND DEVELOPMENTAL EFFECTS OF CHEMICALS

NEED: Current means for evaluating chemical injury to the reproductive function are of limited breadth and uncertain reliability, as are means for assessing chemical injury to the newborn.

RECOMMENDATION FOR PROMISING RESEARCH: Expanded research aimed at understanding mechanisms of chemical injury to the reproductive function.

ANTICIPATED IMPACT OF RESEARCH: Such research is likely to increase the reliability of experimental approaches and assessment techniques for evaluating chemicals that may cause reproductive abnormalities or birth defects.

RATIONALE: Various incidents of the past few years have increased public awareness of the potential for reproductive and developmental problems associated with exposure to hazardous chemicals--for example, the serious birth defects associated with thalidomide use during pregnancy, male sterility resulting from dibromochloropropane (DBCP) exposure in the workplace, vaginal cancer and adenosis in the daughters of women who were treated with diethylstilbestrol (DES) during pregnancy, fetal alcohol syndrome in the offspring of women consuming alcoholic beverages during pregnancy, and increased risk of spontaneous abortion in women who work in surgical operating rooms. As a result, both private and government organizations have been compelled to address potential reproductive and developmental effects of chemical substances to which humans are exposed. But the significance of the effects observed in current reproductive and developmental toxicity tests with respect to humans is unclear; accordingly, the control actions that should be instituted are uncertain.

The development of the ability to determine heritable changes brought about in the genomes of sperm and egg by exposures to toxic substances constitutes a pressing problem, because injury may persist and be expressed in later generations. The Panel believes that advances in molecular biology will help make possible the direct examination of regulation and expression of the human genome and its modification by chemicals.

The Panel has identified several areas of scientific inquiry that may contribute to the development of reliable tests for reproductive and developmental injury. These include research into:

1. Ways in which chemical agents lead to male and female infertility;
2. Morphologic and biologic characterization of sperm as a basis for detecting and understanding the action of environmental agents;
3. Physiology of implantation and the early stages of the maintenance of the fertilized ovum, particularly as related to potential vulnerability to chemical agents;
4. Prediction of the delayed effects of chemical agents on offspring;
5. Relation of postnatal adjustment, growth, and maturation to chemical exposure; and
6. Reproductive ability of groups of humans exposed to certain chemicals, especially at high concentrations.

With the expanded information that would be provided from research on these subjects, it should be possible to develop more specific and more interpretable tests, whose results could shape preventive measures to reduce the likelihood of reproductive and developmental effects of environmental chemicals.

Thus, it is recommended that federally supported research into the mechanisms of chemical induction of developmental and reproductive effects be increased so as to improve the assessment of the potential of environmental chemicals to cause these effects.

IV. DEVELOPMENT OF CELLULAR AND MOLECULAR MARKERS OF EXPOSURE TO CHEMICALS AND MARKERS OF PRECLINICAL EFFECTS

NEED: Sensitive means for estimating individual exposure to chemicals and for detecting early (preclinical) effects are urgently required.
RECOMMENDATION FOR PROMISING RESEARCH: Development and validation of techniques of high sensitivity and specificity for detecting the interaction of chemicals with, and their effects on, tissue components (DNAs, RNAs, hemoglobins, and other proteins) in humans for use in assessment of chemical exposure and risk.
ANTICIPATED IMPACT OF RESEARCH: Such procedures, if fully developed, would be powerful tools for direct evaluation of human exposure to environmental chemicals and its impact on human health.
RATIONALE: Current decision making about potentially hazardous chemicals is handicapped by the lack of anything more than fragmentary data on the levels of human exposure. Current approaches rely on projections derived from exposure models or on data from fixed monitoring stations. Generally, little information is available to measure or verify actual human exposure, especially to individuals.

Specific biologic markers of human exposure (as an adjunct to environmental monitoring) and of their preclinical effects offer a great opportunity to improve the ability to assess the effects of chemicals to which humans may be exposed.

There are several examples of promising research that should be pursued to develop and validate new clinical methods that would enable more accurate description of human exposures and detection of preclinical effects. Such methods should also be validated for animal systems so as to strengthen animal to human extrapolations. Promising methods that should be pursued include:

A. Means for evaluating individual exposure:

1. Analyses of expired air, blood, and urine that can be coupled with tests of toxic activity--for example, the application of short-term mutagenicity tests (e.g., the Salmonella/microsome test) to provide, inexpensively and rapidly, information on a portion of the body burden of mutagens.
2. Analyses for covalent and noncovalent binding of toxic chemicals to proteins or other constituents of assayable fluids or cells to be used as indicators of low-level exposures.
3. Highly sensitive and relatively rapid immunologic and chemical assays for identification of chemically modified nucleic acid bases that are excreted in the urine.
4. Use of monoclonal antibodies to detect specific chemical antigens for rapid screening of potentially exposed populations.

B. Means for detecting early (preclinical) effects:

1. Analyses of chromosomal alterations to detect genetic damage in peripheral blood cells (lymphocytes) of large exposed populations.
2. Detection of toxic effects on the reproductive systems from studies of changes in sperm morphology, including the use of special techniques, such as staining of DNA with fluorescent dyes and the use of rapid-flow cytofluorimetry to detect changes in the DNA organization of the sperm head.
3. Use of recent advances in gene cloning and DNA sequencing to study alterations in the genetic makeup of human cells directly.

With the further development and validation of these indicators, exposure assessments in humans will be strengthened; and these improved assessments will substantially increase the power of epidemiologic studies. Such measures will also improve the basis for extrapolation from laboratory exposures to human exposures. Potential applications include assessment of the consequences of human exposure in industrial and community settings and in connection with toxic-waste dumps. Thus, it is recommended that support be increased for the development and validation of cellular and molecular markers of exposure to chemicals and of markers of preclinical effects of chemicals.

V. DEVELOPMENT AND EVALUATION OF MEANS TO DISTINGUISH AMONG CHEMICAL CANCER-CAUSING AGENTS ON THE BASIS OF MODES OF ACTION

NEED: Cancer induction in humans may often be the result of a complex series of chemical and biologic interactions. Yet, lacking full understanding of these processes, we continue to treat chemical causation of cancer as if it were a one-step process.

RECOMMENDATION FOR PROMISING RESEARCH: Research aimed at refining understanding of the stages of carcinogenesis, including "initiation" and "promotion," and the further exploration of possible fundamental differences, such as those between "genotoxic" and "nongenotoxic" carcinogens.

ANTICIPATED IMPACT OF RESEARCH: Such research may make it possible, in risk assessment and control decisions, to distinguish among potential cancer-causing chemicals on the basis of their modes of action.

RATIONALE: Current evidence indicates that carcinogenesis is a complex process that often involves the induction of a stable and permanent change in the genetic information of a cell (initiation), followed by successive stages that lead to the development of a malignant tumor. A number of factors, including chemical exposures and nutritional status, are known to affect the progression of initiated cells to tumors, either enhancing (promoting) or inhibiting tumor development. If the different modes of action of various chemicals in carcinogenesis were better defined and understood, it might substantially affect the way in which potentially hazardous chemical exposures are evaluated and acted upon.

Too little is known about the importance of man-made and natural environmental agents that may promote or inhibit tumor development. It is evident that man-made chemicals are not the only initiators or modulators of neoplastic changes. Some constituents of natural foodstuffs, uncooked or cooked, can act as initiators, promoters, or inhibitors of tumor formation; such naturally occurring factors may be important contributors and should be intensively explored.

A few chemicals--such as saccharin, TCDD, phenobarbital, PCBs, some chlorinated hydrocarbons, and bile salts--have been shown to induce tumors in animal tissues but show no evidence of genotoxicity; therefore, they may be acting through the promotion of cells previously initiated spontaneously or by unidentified agents. The number of environmental chemicals that may have similar promoting activity is unknown, because of the lack of bioassay systems for their detection. Recent advances in the fields of molecular and cellular biology offer prospects for obtaining information on the mechanisms by which promoting agents exert their effects. An understanding of the mechanisms of action of such chemicals is required for the development of appropriate test systems and the evaluation of their public health significance.

Substantially expanded research that may lead to improved understanding of the significance of promoting agents in human cancer incidence is recommended. Similarly, research should be performed to understand the mode of action of chemicals that may induce cancer without being genotoxic. A wide spectrum of new experimental models should be developed--in vivo models that use animal organs and tissues, cell-culture systems that use human cells and tissue explants, and in vitro systems adequate to permit definition of molecular mechanisms, (e.g., possibly the activation of cellular and viral oncogenes).

The Panel encourages the development of effective short-term tests for the bioassay of environmental chemicals for promoting activity, as an adjunct to those tests that can identify genotoxic initiating agents.

Atherosclerosis is now viewed by some as a proliferative process akin in some respects to cancer. Dietary factors seem to be involved in this process. The exact nature of the mechanisms involved remains to be discovered, but may involve mutation, as well as modulation.

The results of research in these areas are likely to improve our understanding of the differences in modes of action between carcinogens and modulators and may thus form a basis for new approaches to risk estimation and control of chemical exposure.

Thus, it is recommended that a research program be initiated to study differences in modes of action of substances involved in the development of cancer.

VI. DEVELOPMENT OF WAYS TO TAKE THE GREATEST POSSIBLE ADVANTAGE OF EXISTING FEDERAL DATA COLLECTION FOR HUMAN HEALTH EFFECTS ASSESSMENT

NEED: The health impact of environmental chemical exposure of humans is poorly known and largely speculative. Indeed, it may be substantially less than has sometimes been alleged. Reliable estimates require careful epidemiologic studies.

RECOMMENDATION FOR PROMISING RESEARCH: Linkage of existing governmentally collected data sets (such as National Center for Health Statistics, Social Security, and National Institutes of Health data) would provide a useful and efficient basis for epidemiological studies.

ANTICIPATED IMPACT OF RESEARCH: If legitimate concerns over data privacy can be overcome, such linkage would dramatically improve our ability to assess the effects of chemical exposures of humans relatively rapidly and inexpensively.

RATIONALE: There are four events that it is useful to link: exposure, appearance of a preclinical marker, appearance of disease, and death. In the case of many thousands of persons, detailed information has been collected on all four events; but it has been collected at different times, and the sequence has never been linked so as to reveal important patterns. Cost-effective research could be undertaken if the obstacles to such linkage could be removed. For example, within a single federal agency--the National Center for Health Statistics (NCHS)--it would be possible to:

- Identify relationships among occupation, lifestyle, and many environmental characteristics to specific causes of death by linking information obtained on the National Health Interview Survey to deaths identified through the National Death Index (NDI);
- Identify preclinical markers of serious disease by linking information obtained in the Health and Nutrition Examination Survey

(HANES) to deaths identified through the NDI or, even better, by re-including some persons in HANES later; and

- Identify preconceptional and pregnancy determinants of disease by linking data from samples of births to deaths of infants identified through the NDI.

The possibilities are further expanded when one considers eventual linkages between data of NCHS and those collected by other federal agencies. Two come quickly to mind: Linkage between NCHS and the Social Security Administration (SSA) to link the SSA 1% sample of continuous work histories with deaths ascertained through the NDI, and linkage between NCHS and the National Institutes of Health (NIH) data. NIH, in the course of many very large case-control studies and even larger experimental-intervention studies, has assembled detailed environmental, social, and medical information on many thousands of normal persons (i.e., the controls in the case-control studies, the nonexperimental group in the intervention studies); those data should be followed through the NDI with causes of death linked to the information obtained in the NIH research.

These are only a few of many possibilities. Funding, although needed, is not now the principal constraint. The concern for privacy severely constrains efforts to link data. An approach needs to be developed to determine and protect important aspects of confidentiality--but only those aspects, because the linkages proposed here can provide information of great value to society.

Thus, it is recommended that the policy issues related to data linkage be addressed and that appropriate linkages be made so that existing federal data collection activities can be used to the maximal extent possible to assess the potential human health effects of chemical exposures.

REPORT OF THE RESEARCH BRIEFING PANEL ON MATERIALS SCIENCE

Research Briefing Panel on Materials Science

Theodore H. Geballe (Chairman), Stanford University
Martin Blume, Brookhaven National Laboratory
Albert M. Clogston, Los Alamos Scientific Laboratory
Billy Lee Crowder, IBM
Arthur C. Damask, Queens College of CUNY, Flushing, New York
Henry Ehrenreich, Harvard University
Fred Gamble, Exxon Research & Engineering Company
Stig Hagstrom, Xerox Corporation
John K. Hulm, Westinghouse Research Laboratory
Herbert Johnson, Cornell University
Venkatesh Narayanamurti, Bell Telephone Laboratories
William D. Nix, Stanford University
Donald R. Uhlmann, Massachusetts Institute of Technology
John Wachtman, Jr., National Bureau of Standards

Invited Guests

Malcolm R. Beasley, Stanford University
Robert L. Byer, Stanford University
Robert A. Huggins, Stanford University
Lewis H. Nosanow, National Science Foundation

Staff

Klaus Zwilsky, Executive Director, National Materials Advisory Board
Donald G. Groves, Staff Scientist, National Materials Advisory Board
Raymond P. Girouard, Special Assistant (Stanford University)
Allan R. Hoffman, Executive Director, Committee on Science, Engineering, and Public Policy

REPORT OF THE RESEARCH BRIEFING PANEL ON MATERIALS SCIENCE

ABSTRACT

Opportunities exist for extending the frontiers of materials science and for increasing the flow of materials that can lead to new and improved technologies. Materials with the potential for higher performance because of their improved strength, corrosion resistance, and reduced energy consumption are expected to be developed from new types of composites, structural ceramics, and polymers. Materials with the potential for high technology with improved electrical, magnetic, and optical properties are expected from new semiconductor composites, superconducting composites, and optical crystals, glasses, and fibers. Nonequilibrium materials obtained by rapid quenching, surface implantation, and other means are expected to provide a reservoir of technology for reducing problems such as the U.S. dependence on strategic materials. Cooperative phenomena such as solid-liquid transitions, magnetism, superconductivity, metal-insulator transitions, sound waves, charge density waves, and spin waves can be better understood both because of theoretical advances and because of the availability of new materials. The precise description of clean surfaces now available can be extended to technologically important surfaces and provide direction for producing urgently needed new catalytic materials. The promise of developing the science of fracture now based on macroscopic concepts down into the atomic regime and to be able to make quantitative predictions may be realized. The opportunity to focus on microscale phenomena permeates almost all materials science. The time is ripe for new tools and concepts for exploring phenomena and characterizing materials on a microscale. Instrumentation is an essential ingredient; there is an opportunity to enhance progress in almost all of the areas covered in this report by making state-of-the-art instrumentation and synthesis techniques available to more students and practitioners and by encouraging further development of new methods of synthesis and characterization.

In summary, many areas of materials science offer exciting opportunities for advancing basic science and technology. The Panel is recommending six that seem most promising and timely.

1. Increasing access to the state-of-the-art equipment that is available commercially.
2. Discovering new materials using techniques for synthesis and characterization that depend on physics that is understood but has not been utilized as yet.
3. Making more support and modern equipment available for graduate student training.
4. Investigating phenomena and structures in systems that contain "small" numbers of atoms such as surfaces, interfaces, and grain boundaries.
5. Extending the scientific understanding of the basic science that pervades processing techniques.
6. Developing a science of fracture and other modes of failure based on atomic level models.

INTRODUCTION

Materials science is a broad interdisciplinary research field that brings together the ideas and theoretical understanding and experimental skills of physics, chemistry, metallurgy, and ceramic science. New materials of scientific and technological interest are synthesized in which the mechanical, thermal, chemical, electrical, optical, magnetic, and other physical properties are controlled by understanding the role of defects, impurities, composition, and structure. Figure 1 illustrates the key role that materials science plays in technology transfer. Among the many past accomplishments of this type of interdisciplinary research is dislocation-free silicon--the building block of silicon-chip technology. An important activity within materials science is the exploration of new materials that may have unusual or unexpected combinations of properties.

Theory and experiment in condensed matter studies are converging as condensed matter theory becomes more proficient at dealing with model systems that approximate real materials in essential ways and as methods of preparing and characterizing materials become more precise. Materials science is well on the way to establishing a fundamental basis for the technology of new engineering materials. The need to optimize more than one property (see Figure 2) frequently calls for conflicting methods of processing. Establishing acceptable trade-offs is a function that benefits from an understanding of the basic phenomena. The development of its constituent fields to their present level of sophistication has been brought about both by government and industrial support but will increasingly be a major responsibility of government because of the long lead times involved in bringing the basic technology of the complex new materials to commercial viability.

Strategic materials, such as chromium, cobalt, and platinum are largely obtained from abroad, and it is urgent that substitutes for them be found. The most promising and perhaps only nonpolitical route by which alternatives to such strategic materials can be found is through materials science. In general, it can be said that materials are strategic because of the unique properties they provide, such as

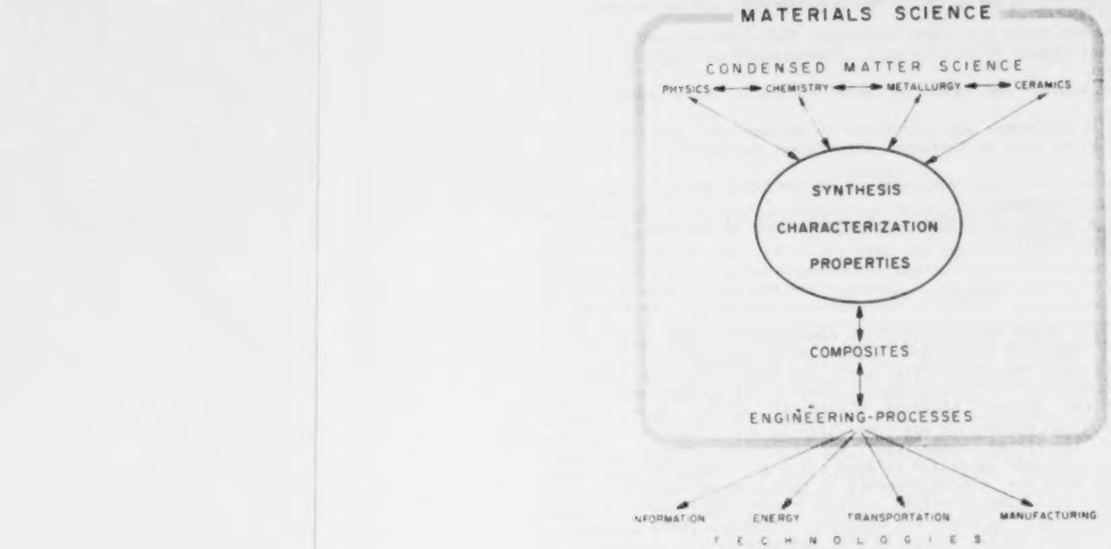

MATERIALS SCIENCE
CONDENSED MATTER SCIENCE
PHYSICS
CHEMISTRY
METALLURGY
CERAMICS
SYNTHESIS
CHARACTERIZATION
PROPERTIES
COMPOSITES
ENGINEERING-PROCESSES
INFORMATION
ENERGY
TRANSPORTATION
MANUFACTURING
TECHNOLOGIES

FIGURE 1 Materials science and engineering occupy a key role in the flow of condensed matter science into the industrial sector.

corrosion resistance, hardness, strength, and oxidation resistance at high temperature and catalytic activity. Higher-priced substitutes exist for some applications, but in others any substitution of present materials results in a loss of performance. However, trends in the development of new high-performance materials promise both cost-effective and/or performance-effective substitution. For example, polymer-matrix materials, glass, and glass-ceramic coatings are promising for corrosion-resistant chemical processing. Surface-alloyed materials offer improved hardness and corrosion resistance. Structural ceramics offer wear resistance and high-temperature service capabilities. Rapidly solidified crystalline metals offer new families of alloys with extended solid solutions, some of which have high strength, hardness, and high-temperature capability. Unpredicted discoveries of new materials can reasonably be expected to occur, and some entirely new solutions to strategic materials problems may appear as a result.

It should be noted from the above considerations that recommendations for research support cannot be simply product oriented. The

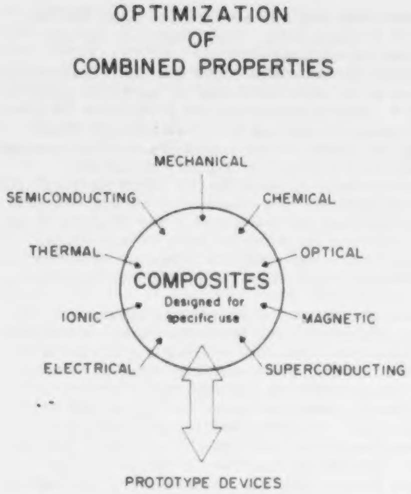

OPTIMIZATION
OF
COMBINED PROPERTIES
MECHANICAL
CHEMICAL
OPTICAL
MAGNETIC
SUPERCONDUCTING
ELECTRICAL
IONIC
THERMAL
SEMICONDUCTING
COMPOSITES
Designed for specific use
PROTOTYPE DEVICES

FIGURE 2 The multidisciplinary background of materials science is frequently brought into play when it is necessary to optimize more than one property.

fundamental approaches to materials science provide the basis for a wide range of applications, for example, particle-solid interaction studies have led to new ion implantation instruments by which accurate diffusion profiles can be measured and semiconductor systems can be made. Without the basic studies, such tools and instruments would not exist--the technology would not exist. New and developing technologies contribute most to the growth of our nation's industrial strength as well as to the generation of new types of employment.

The panel has addressed the opportunities for high-leverage scientific investment in materials science using the following procedure. In this report, classes of materials are identified as high technology for utilization in electronic, optical, magnetic, and acoustical devices and as high performance for utilization in mechanical, chemical, or thermal devices. The fundamental properties of these materials are described and the basic science required for increasing the understanding discussed. Underlying the research is the need for the accurate characterization of materials. Opportunities for new instrument development are illustrated.

Materials science, with its multiple roots and many-branched interfaces, is less monolithic than, say, astronomy or particle physics; consequently, obtaining a consensus of research priorities is less straightforward. Much research involves relatively small numbers of investigators. As a result, there can be rapid changes in response to new stimuli so that worked-out fields are rapidly abandoned.

Choices must be made between emphasis on the large national facilities versus small institutional-based research. Both are essential, but the balance must be continuously monitored. The ability to develop and bring large-scale facilities to fruition must be maintained, even at the expense of phasing out or reducing the scope of still-valuable facilities that are no longer competitive in producing exciting new science.

Four national synchrotron radiation-source facilities are now in operation and serve many different constituencies in the scientific, medical, and industrial sectors. The new science and advances in instrumentation, characterization, and technology that have been accomplished, and those that should be forthcoming, justify, at present, financial support at something close to the projected levels.

Neutron spectroscopic and radiation facilities exist at national laboratories (Brookhaven, Oak Ridge, Los Alamos, Argonne) and other government laboratories. The problem of how to set priorities and take advantage of new advances has already been addressed by the Review Panel on Neutron Scattering (Brinkman) committee. The cost-effectiveness of national facilities such as the Magnet Laboratory at MIT, the submicron facility at Cornell, the Solar Energy Research Institute, and other special-purpose laboratories should be evaluated with respect to the totality of research and technology being done by smaller institution-based groups. In general it is believed that the most effective use of resources is to build on the expertise at existing laboratories rather than to create new ones. The training of workers with state-of-the-art equipment of new research must be an important consideration in order to overcome the present-day shortage of good materials scientists in industry.

HIGH-PERFORMANCE STRUCTURAL AND SPECIAL-PURPOSE MATERIALS

These are materials whose intrinsic characteristics, which are determined by their components and the methods by which they are synthesized, enable them to give improved performance.

Composites and Structural Ceramics

Use of improved composites and structural ceramics can increase energy efficiency; reduce dependence on strategic materials; and help keep U.S.

products such as aircraft, automobiles, and power turbines competitive with foreign ones.

Promising composites include high-modulus fibers (such as graphite or silicon carbide) in several types of matrices, including polymers, metals, and ceramics. Higher strength-to-weight ratios, high toughness, improved fatigue resistance, and high corrosion resistance can be achieved. Realization depends on improved interfacial characterization and control and on processing science to achieve optimum distribution of the components. Advances in spectroscopy applied to interface characterization and in the rheology of fiber/liquid suspensions give promise of providing a science base to support processing developments.

Promising structural ceramics include silicon carbide, silicon nitride, and related materials. Sufficiently high levels of strength and toughness already achieved permit short-term (100-hour) application up to 1100°C. Applications requiring a longer service life and higher temperatures appear possible if stronger grain boundaries can be achieved. Use of these improved materials can increase energy efficiency in diesel and turbine engines and extend service life in many types of machinery that are subject to high wear rates. Grain boundaries in the present material have relatively poor long-term strength and oxidation resistance because of additives needed for sintering. Improved powder production and control of mass transport in the boundaries appears possible. Modern surface characterization methods and analytical electron microscopy offer the opportunity to place the design of grain boundaries (including impurity concentration effects) on a scientific basis to guide optimization of sintering with reduced levels of long-term creep and oxidation.

Another important class of new structural ceramics is the family of transformation-toughened ceramics typified by alumina strengthened by zirconia inclusions. Significant increases in strength and toughness have been achieved. The technique is promising for strengthening other ceramics, but fundamental scientific questions concerning the mechanism of strengthening need to be answered. Several theoretical models exist; detailed experiments involving structural characterization on a microscale are required and appear possible.

Polymers

The use of polymers has increased by 10 to 15 percent each year in the last 50 years--the fastest growth rate of any class of materials. Polymers are now used, or are considered for use, as critical components in materials systems ranging from structural composites and ophthalmic lenses to semiconductor packaging and chip fabrication.

Among the areas where future developments are limited by a lack of scientific understanding, and where improved understanding could have great impact on technology, four are particularly noteworthy: (1) liquid-crystal polymers, (2) high-temperature polymers, (3) lithographic polymers, and (4) polymeric adhesives.

Liquid-crystal (rigid-chain) polymers offer exceptional, and previously unobtainable, combinations of properties (a tensile strength approaching that of steel yet with a density less than 1.5 g/cm^{-3} and relatively high thermal conductivity and low thermal expansion) whose applications range from reinforcing fibers to bulk moldable materials, from aircraft with larger payloads to automobiles with greater fuel efficiency. The effective development of this relatively unexplored class of polymers depends critically, however, on acquiring improved understanding of the statistical physics of rigid-chain molecules and of the structure, rheology, crystallization, and relaxation behavior of the materials.

High-temperature polymers (polymers with use temperatures exceeding 300°C) offer outstanding opportunity in areas ranging from novel protective coatings to semiconductor packaging. The achievement of these opportunities depends on an imaginative effort in synthesis, processing, and characterization. Promising chemical approaches such as combining organometallic catalysts with polymer substrates are largely unexplored, as are the chemical and structural changes that take place on curing and the effects of these changes on mechanical, thermal, and chemical properties.

Lithographic polymers are essential elements in the fabrication of integrated circuit devices. The demands of large-scale integration and production efficiency require new polymers with increased sensitivity, resolution, and difference in etch rates between exposed and unexposed regions. These demands require a better understanding of ultraviolet, x-ray, and electron-beam exposure and, in turn, of polymers with the requisite sensitivity to these radiations. Improved polymers in this area would provide great leverage for integrated circuit manufacturers; yet only small volumes of materials are required, and this provides reduced incentive for industrial groups with traditional polymer expertise. Future developments in lithographic polymers will depend on fundamental investigations of the interaction of various radiations with polymers, synthesis of novel polymers with tailored combinations of properties, novel approaches to the resist function (as polymers that fall apart and vaporize on exposure to appropriate radiation), and understanding of relevant processes such as reactive ion etching.

Novel polymeric adhesives offer the promise of effecting radical changes in manufacturing technology and product design. At the same time, developments of the past decade in surface spectroscopic techniques and fracture mechanics offer the possibility of obtaining critical new insight into the nature of adhesion, its sensitivity to environmental factors, the fracture mechanics of adhesive joints, and the role of coupling agents and interphase material. Work in these areas, dealing with polymers whose chemical and physical structure is tailor-made, should dramatically change current technology. The need for understanding is critical; the tools for providing that understanding already exist.

High-performance materials are frequently subject to high material stresses in service. The failure of such materials can be quite costly or even catastrophic. Much effort and cost go into ensuring safe and reliable service. Opportunities exist to increase reliability and reduce costs by coupling improved understanding of failure processes with improved diagnostic techniques.

HIGH-TECHNOLOGY ELECTRICAL, MAGNETIC, ELECTRONIC, AND OPTICAL MATERIALS

These materials are the basic building blocks of the electric power, control, computer, and communications industries. Advances in electronics, in particular microcircuits, have been so rapid in the past two decades that they are producing a major revolution in information technology. These advances are based on materials-processing technology at a sophisticated level where complex microcircuit structures composed of many layers of different thin films can be fabricated with precise control on a scale of 1 micron or less. The materials technologies discussed below are of growing importance.

Electrical and Magnetic Materials

The power generation, distribution, and control industries already rely heavily on sophisticated magnetic alloys and advanced composite polymer insulators. New materials that are of growing importance in electrical applications include large area silicon crystals for power electrical devices, which play a rapidly growing role in AC to DC power conversion and variable frequency devices. Advanced materials under development include high-field high-current superconducting composites for making electrical generators more efficient. Improved multifilamentary superconducting composites are essential for providing magnetic confinement in nuclear fusion plants. High saturation, low-loss amorphous magnetic materials have technological promise for transformers and machines.

Electronic Materials

Beginning with basic work at Bell Telephone Laboratories on the junction transistor in 1948, a new materials-based solid-state electronic technology has developed in the past 30 years that has produced staggering advances in the arts of communication, computers, command, and control. The first wave of this new technology was based on the ability to grow defect-free, high-purity single crystals of silicon and the ability to fabricate micron-scale devices with metal-oxide-semiconductor layers arranged in micro-scale transistor circuits. The fabrication of microcircuits represents probably the most advanced materials processing developed to date. The evidence suggests that we have only scratched the surface of the opportunities in this area and that enormous growth is still to come. The new electronic technologies now emerging will have major influence on the entire spectrum of U.S. industry--in communications, automation, command, and control.

These new technologies will utilize not only silicon but also semiconductor crystals such as gallium arsenide, indium phosphide, mercury cadmium telluride, and silicon carbide. Integrated circuits will be coupled with advanced sensors to detect signals over the entire electromagnetic spectrum. Signals will be processed at higher and higher speeds and will involve cryogenic semiconductor and superconducting

Josephson-junction thin-layer devices. They will use other, nonelectromagnetic methods such as surface acoustic wave techniques. Each of these requires a multitude of new electronic crystals, new thin-film and multiple-layer materials and processing methods using photons, coherent photons, electron beams, x-ray, and other advanced processing techniques. Magnetic films with vertical anisotropy are generating excitement because of their greatly increased digital storage capacity. Amorphous rare earth-based materials may be the long-sought medium for magnetooptical recording.

Optical Materials

Running close to the electronic revolution is a new era of optical technology that ultimately will be the main channel for most information transmission in the world. This will involve national broadcast, cable and telephone signals, computer networks, and control circuits in industrial plants. This new communication medium is being made possible by three independent but converging materials developments:

1. Low loss optical fibers.
2. Lasers, in particular micro-sized solid-state lasers.
3. Solid-state photon detectors and electrooptical devices.

Rapid advances in optical materials and electrooptical devices are expected in the next 20 years, which may lead to other revolutionary changes in process control. A plethora of new electrooptical, acousto-optical, and other cross-coupled materials are emerging, such as (optical) damage-resistant lithium niobate and thallium arsenic selenide. Single crystals will be required in most cases and will require new advanced methods of growing high-purity, defect-free single crystals, such as single-crystal fibers, which can be used for remote sensing. This general field will greatly influence avionics and defense technology and represents one of the greatest U.S. technological opportunities on the horizon.

Advances in glass (Nd-doped) and crystal (garnet) laser materials can be expected to double and even triple the efficiency of the lasers in the next 10 years. Even more important than increased efficiency might be the reduced size and weight that will permit the mounting of multikilowatt lasers on robot arms for such applications as cutting, drilling, and welding.

NONEQUILIBRIUM MATERIALS

Materials far from equilibrium are typically produced by processes that prevent lower free-energy or ground-state structures and compositions from forming. A particularly valuable method for controlling the kinetics is by quenching from the melt and the vapor. A remarkable variety of nonequilibrium structures can be produced, with resulting properties that are of great scientific and technological interest, by varying the processing conditions. Opportunities exist in both the understanding and the control of the material transformations that occur in these processes. Spectroscopic techniques for very rapid measurement of local temperatures, reaction rates, and mass transport of chemical species, including short-lived intermediate species, can assist in building a science base for processing control and for suggesting process modification. Opportunities exist for improving the characterization of the novel structures produced. These include new glasses, extended ranges of solid solutions, gradient structures, layer structures, and modified surfaces. An emerging field of materials science concerns the relation of these novel structures to unique and desirable properties on the one hand and to the process details required for producing these structures on the other.

The existence of new kinds of materials and processes may provide a way to reduce our nation's dependence on certain strategic elements. For example, rapid solidification methods provide ways of creating alloys with entirely new chemical compositions, compositions that could not be explored previously because of the limitations of conventional processing. The addition of transition elements such as Fe and Mn to aluminum alloys for increasing elevated temperature strength is now possible because of this development. Thus, a much broader range of opportunities now exists for selecting more abundant elements for the synthesis of high-performance materials.

Processing of surfaces by laser heating can be used to improve the wear and corrosion resistance of metals significantly. The strategic material chromium, which is required for corrosion resistance in stainless steels, is needed primarily at the surface where the corrosion takes place. Through the use of new surface processing methods it might be possible to create stainless steels using but a small fraction of the chromium now required for this application. Ion implantation opens the way to forming new surface alloys. These can be crack resistant, corrosion resistant, and wear resistant. Other alloys might have catalytic properties.

Substitutional nitrogen, an important impurity for the basic understanding of deep-level defects in semiconductors, was first produced in silicon by ion implantation and laser annealing after other techniques failed. This technique has recently been used to enhance by a factor of 10^5 the photosensitivity of transparent ferroelectric ceramics and to cause significant changes in the spectral sensitivity. These new materials find use as nonvolatile and electrically erasable image-storage devices.

OPPORTUNITIES IN CONDENSED MATTER SCIENCE

Nonequilibrium structures provide a powerful approach to the testing of our basic understanding of condensed matter. The extended compositions available allow theories of physical, chemical, and in some instances mechanical properties to be evaluated over much wider ranges of the relevant variables. For example, important yet subtle concepts of the role of structural (atomic) order/disorder may be tested through studies of amorphous and crystalline materials of the same chemical composition.

Solid-State Cooperative Phenomena

In a fundamental sense, all solids and liquids exhibit properties based on the interactions of very large numbers of electrons and atomic nuclei whether these are best represented as neutral atoms, charged ions, or mobile electrons moving within a lattice of fixed positive ions. The range of phenomena encountered in such complex, many-body systems is great. In particular, such systems exhibit emergent, cooperative phenomena that are not present in disassembled arrays of separate atoms. The simplest such phenomenon is the crystallization of solids into regular arrays of atoms or ions. Others are the thermodynamic transitions between different solid and liquid phases, magnetism, superconductivity, metal-insulator transitions, sound waves, charge waves, and spin waves. As mentioned below, some particularly important cooperative phenomena occur in materials systems that are effectively one- and two-dimensional. Since the site of action of many chemical and electronic phenomena, such as catalysis and transistor action, is on surfaces, the discovery of cooperative phenomena in two-dimensional systems opens exciting scientific and technological opportunities.

The difficulties of dealing theoretically and experimentally with cooperative phenomena in condensed materials led physicists at first to seek out the simplest systems that might be understood on an atomic basis, such as the pure monovalent metals and pure single crystal semiconductors. In these cases useful approximations could be made based on the motion of single electrons, and cooperative phenomena did not play a prominent role. At this earlier time, a fundamental atomic understanding of the stabilization of crystal phase, magnetism, superconductivity, and other cooperative phenomena was beyond the explanatory power of condensed matter physics.

Today, because of the development of more powerful theoretical and experimental methods, the situation is quite different. The ground state energy of ordered crystal phases can be calculated in many cases from first principles to high numerical accuracy using methods that take account of the many-body interactions. Matter is studied under extreme conditions of high pressures and low temperatures where new thermodynamic states and transitions can be found to test new ideas. Second-order phase changes are well understood because of the introduction of renormalization group methods. Superconductivity is better understood on an atomic basis as a result of a series of theoretical developments starting with the Bardeen-Cooper-Schrieffer (BCS) theory but is not yet capable of predicting transition temperatures or explaining why critical temperatures above 23 K have not been found experimentally. Magnetism is now much better understood both in terms of the atomic interactions giving rise to spin alignment and the phase transitions between magnetic and nonmagnetic states. Two other cooperative phenomena closely related to magnetism are now beginning to be well understood. The first is valence fluctuation whereby rare earth atoms in certain compounds do not possess a single fixed valence but fluctuate dynamically between two charge states. The second is the metal-insulator transitions in which transition metal oxide and some other materials exhibit a transition temperature at which the compound changes its conductivity by many orders of magnitude from a conducting to an insulating state. Both of these phenomena share with magnetism the fact that they are bulk properties of matter and therefore can have large macroscopic effects. They have not yet found a firm place in materials technology but deserve further attention because of their potential for useful applications.

It is a fact of quantum physics that systems exist in nature that effectively exhibit one- and two-dimensional behavior and have many practical effects. The basic reason is that a limitation on the transverse motion of a quantum mechanical system, because of zero point effects, makes the system effectively one- or two-dimensional for small enough excitations. The most exciting recent developments in cooperative phenomena have been in reduced dimensionality systems. In two dimensions, for example, strong cooperative effects have been found in semiconductor inversion layers. In the presence of a magnetic field, two-dimensional inversion layers on Si and GaAs show a totally unexpected, accurately quantized Hall effect that is leading to an improved determination of the ratio of the fundamental constants e/h. In a two-dimensional system comprised of electrons trapped on the surface of liquid helium, Wigner crystallization has been observed for the first time, whereby a charged plasma of electrons condenses into a two-dimensional electron crystal. In other two-dimensional systems, such as monolayer liquid crystal films and layers of atoms intercalated into layered transition metal disulphides, the melting of two-dimensional crystals has been observed.

Another partially understood cooperative phenomenon is that of localization of electron motion in disordered solids. This is an effect whereby an electron of sufficiently low kinetic energy may actually become trapped in a localized region of a disordered conductor. The effect, particularly marked in one- and two-dimensional systems, leads to strong departures from ohmic conduction in partial agreement with theory. This effect could become one of practical significance in large-scale silicon integrated circuits as the characteristic dimensions of conduction paths shrink below 0.1 micron, particularly if integrated circuits of the future are operated at cryogenic temperatures to increase electron mobility and improve heat transfer.

Another class of one-dimensional cooperative phenomena arises in the conducting polymers such as polyacetylene where conduction takes place by electron transfer along linear chains of molecules or through linear stacks of molecules such as in the organic conductor tetrathiafulvalene-tetracyanoquinodimethane (TTF TCNQ). These materials show many cooperative phenomena such as spin waves, charge waves, and the so-called Peierls instability. In some cases they show soliton formation whereby a coupled electron-nuclear system develops a high amplitude motion that propagates as a charged entity with many of the properties of a free noninteracting particle.

On the very forefront of the physics of cooperative phenomena is a poor but growing understanding of the way in which nonlinear effects in systems far from equilibrium showing charge or mass transport can lead to bifurcation of their periodic response followed by the onset of chaos and turbulence as the driving field is increased. Examples of such instabilities occur in hydrodynamic flow, chemical reactions, phase transitions, and nonlinear electronic circuits.

Cooperative phenomena, as the above examples show, generally tend to arise, or find their strongest expression, in the more complex material systems such as high-temperature superconductors (the A15 structures), bubble memory ferromagnets (the yttrium iron garnets and analogs), the quantized Hall effect (artificially layered GaAs), the metal-insulator transition (vanadium oxides), two-dimensional melting (liquid crystals, monolayers of rare gases), and soliton formation (polyacetylene). Research efforts form strong links between condensed matter physics, metallurgy, and solid-state chemistry and lead to the convergence of these fields into a unified materials science.

Surface Science

The study of clean surfaces is an active and rapidly progressing field that moves hand in hand with ultra-high vacuum technology. The structure of the different faces of metals and semiconductor crystals can be determined in great detail both theoretically and experimentally. The calculated displacements of the outer atoms from their unrelaxed position (i.e., a position determined by an integral number of lattice translations from the interior) are often in good agreement with the reconstructed surface symmetry observed by low-energy electron diffraction and with the atomic positions determined by angular-resolved photoelectron spectroscopy and other surface-sensitive techniques. The positions and bonding of ad-atoms can also be calculated and observed. Field ion microscopy can be used to follow the motion of the ad-atoms. There is an opportunity for extending the studies to many more surfaces, including technologically important ones, to learn more about the behavior of ad-atoms and to move in the direction of being able to understand the atomic behavior of surfaces involved in chemical reactions.

Surface chemistry includes such areas as catalysis, in which chemical reactions take place on surfaces having highly controlled chemical and structural properties; corrosion, in which undesirable modifications of materials properties are generated by surface chemical reactions; and photochemical surface reactions, which offer promise of new ways to synthesize chemicals.

The phenomena involved in corrosion are of technological importance in areas ranging from buildings and boats to nuclear waste disposal and high-performance batteries. Recent advances in surface spectroscopic techniques have opened the possibility of obtaining new insight into the underlying chemistry and physics and the opportunity of developing more detailed models to describe corrosion phenomena. The role of corrosion inhibitors can also be explored in depth, and this in turn should lead to descriptions of the relevant thermodynamics and kinetics.

The area of adhesion provides another example of surface-dominated phenomena. The ability to form adhesive bonds that are stable in hostile environments would be greatly facilitated by improved understanding of the chemistry, structure, and mechanics of interfaces. Advances in surface spectroscopic techniques and in techniques for chemical modification of surfaces offer outstanding opportunities for developing new insight.

Heterogeneous catalysis is a field based on solid materials whose surface chemistry enables them to facilitate the transformation of less valuable molecules into more valuable ones. Heterogeneous catalysts are at the heart of the petroleum, petrochemical, and chemical industries. As such, they are usually thought to lie in the domain of chemical engineering and are not considered a part of materials science. This state of affairs, while understandable, is not a reflection of the nature of the situation. There is an opportunity for extending surface characterization techniques to the study of the kinetics of surface-catalyzed reactions.

In the petroleum and petrochemical area, the shift is to feedstocks that are heavier (larger molecules) and dirtier (lots of S, N, etc.). Crude oil quality is decreasing. Improved catalysts are needed to reduce the size of molecules as well as their S and N contents. In addition, while the synthetic fuels era is further off than was thought of few years ago, it is only delayed 10 years or so from earlier scenarios. This extra breathing time can be used to improve our technology. For example, shale oil contains a higher concentration of nitrogen than ordinary petroleum. This must be removed. Catalysts are sought that will permit the removal of nitrogen selectively, i.e., without the consumption of excessive hydrogen.

One of the largest technologically important surface problems is associated with the internal surfaces of the porous rocks from which we extract oil. Oil reservoirs are formations of porous rock of enormous surface area filled with oil. When oil is produced, only a fraction of it, less than half, comes from the rock. The balance is held in place, absorbed on the rock's internal surfaces. Thus, we have in place in the ground in abandoned fields oil "reserves" that exceed all that have been consumed to date.

An underutilized energy resource is natural gas in remote locations. Economic processes need to be developed that can convert methane into liquids. A catalyst that could convert methane directly to such products does not exist. If it did, the economics for such a process would improve dramatically.

Syn-gas or CO/H_2 is going to be a major feedstock of the future. It will be the product of coal gasification and from it much of our fuels and petrochemicals may be derived. Today catalysts do not exist that can take CO/H_2 selectively to desired products. Instead, a spectrum that must be separated and processed further is obtained.

Strength and Fracture Phenomena

Ultimate strengths of materials are known from theory to be very high and have actually been observed to be so in a few defect-free, structurally perfect materials. There is a clear need for further study of the fracture physics of the newer high-strength materials, especially fibers such as graphite and high modulus polymers such as Kevlar. The fundamental factors controlling fracture of these materials, and their relation to process parameters, environmental conditions, etc., are not well understood.

Failure of most practical materials occurs at much lower stresses, because of a multiplicity of factors, including stress concentrations, the motion and interaction of defects, and environmental factors. The role of defects has been conceptually clear for many years, and some progress has been made in treating the behavior of individual dislocations, cracks, and vacancies. The latter are of special importance in high-temperature fracture.

The extreme mathematical complexity of dealing with the interacting systems of defects involved in deformation and fracture has largely prevented the application of fundamental theory to practical conditions. The situation is now changing as a result of recent studies on both the atomic and continuum levels. On the atomistic level, modern computational methods have made it possible to treat defect interactions and to extend such models to large aggregates of defects.

Several recent developments offer the promise of leading to a true science of fracture based on atomic-level models. A realistic two-dimensional model of the atomic structure of a crack has finally been solved in closed form and used to describe a form of slow fracture associated with "lattice trapping." At the same time new developments in continuum descriptions of the near-crack region are being used in experimental studies of crazing in polymers. On yet another front, a detailed chemical model has been used to predict which polar molecules are potent stress-corrosion agents for silicates.

The way now seems open to combine discrete lattice effects, interacting defects, molecular models of chemical reactivity, and continuous mechanism treatments in a stress field to produce a fundamental theory of fracture that could be quantitatively tested against experiment. A relatively modest increase in support might well have high leverage.

SCIENTIFIC BASIS OF SYNTHESIS AND PROCESSING

During the last 20 years there has been a rapid growth of synthesis methods performed in a vacuum. Examples of such techniques are evaporation, sputtering, chemical vapor deposition (CVD), and plasma deposition. For many of these techniques the basic physical mechanisms behind the deposition are still not well understood--for example, the sputtering process or the plasma deposition. Two techniques look particularly promising for the epitaxial growth of semiconductor materials with a much larger field of applications. These are molecular beam epitaxy (MBE) and metallo-organic chemical vapor deposition (MO-CVD).

The development of ultra-high vacuum (UHV) technology ($<10^{-10}$ torr) has allowed the development of MBE as a method for the synthesis of structures by controlling the deposition of each layer as it is being deposited. The broader application of UHV technology to materials synthesis will allow fabrication of high-purity, better-controlled materials. This approach is thought to be far more interesting and promising than materials synthesis under reduced gravitational gradients (space synthesis).

The other epitaxy growth technique that seems particularly interesting is MO-CVD. Here the growth occurs by the decomposition of metallo-organic molecules in a thermal reactor. The technique allows a close control of the growth by monitoring the gas flows into the reaction chamber. It provides exceptionally high quality semiconductor materials, especially in the III-IV compounds family.

New process development is often a quasi-empirical activity. Processes are inherently complex; the important parameters are usually numerous and sometimes not well defined. The scientific base for the various stages of a "typical process" is often inadequate. Consequently, process development may be a frustrating and time-consuming activity.

The science of processing would be greatly advanced by increased understanding of some basic phenomena that pervade many processing techniques. Nucleation is an important example. Nucleation of a new phase is an important aspect of virtually all processing techniques, microscopic and macroscopic. Processes that would benefit from a better understanding of nucleation include rapid solidification, vapor deposition, production of ultra-fine and monodisperse powders, and sol/gel methods. Lack of understanding of transport across and along interfaces is another area that limits process development. New studies of mass and heat transport across interfaces, often with unconventional geometries and boundary conditions, would facilitate process development. In particular we note the opportunities in ion beam technologies with the use of ion accelerators for both materials characterization and materials synthesis/modification (ion implantation) and in inelastic scattering of neutral atoms.

Processes based on sol/gel technology offer promise for the production of novel glasses and ceramics--materials with novel microstructures and properties. Developments in this area depend on improved understanding of the transitions: sol-gel-dried, gel-glass-crystalline ceramic, of the conditions necessary for the formation of gels (including the effects of pH and anions), and of the structure and properties of gel-derived materials. Also of promise are mixed organic-inorganic materials derived from gels, although not enough is known about their structure and properties.

Techniques for producing monodisperse fine powders depend on improved understanding of nucleation kinetics (in the gas phase as well as in condensed phases), of the growth and coalescence kinetics of small (500-5000 Å) particles, of the interactions between such particles, and of the phenomena involved in comminution.

Conversion of particles or dense bodies depends on improved understanding of particle-binder interactions, of drying phenomena, of sintering kinetics (particularly the sintering kinetics of covalently bonded materials), and of the role of dopants as sintering aids or grain growth inhibitors. Exceptional processes for enhanced dimensional control include injection molding of ceramics and casting of polymer-matrix composites that display no volume change on cure. Developments in this area depend on improved understanding of the rheology of highly loaded suspensions, of the dynamics of injection molding for such suspensions, and of the dependence of volume change during polymerization on chemical structure.

There is a great need for fundamental studies to elucidate the chemical and physical structures of polymeric materials. Controversy exists concerning the basic microstructural units of semicrystalline polymers (adjacent re-entry folding versus switchboard models for the lamellar plates); reptation and (snake wiggling) center-of-mass diffusion are both strongly advocated as the dominant processes for diffusive transport in the liquid state; important structural features of liquid crystal polymers remain largely uncharacterized; the chemical and structural characteristics of thermosetting polymers, and their dependence on process history, are substantially unexplored; and the detailed roles of such modifiers as rubbery second-phase material in semi-crystalline polymers have not been clarified. Work in areas such as these should yield improvements in our understanding of and our ability to modify the properties of the materials.

INSTRUMENTATION

Access to modern instrumentation is a prerequisite for progress in nearly all areas of materials. In the cycle, new materials synthesis/processing/characterization/application, the development and use of new instruments has allowed a shortening of the cycle time and the development of completely new classes of materials. The needs of the materials scientist in this cycle have formed an important impetus for the development of new and sophisticated instruments. Discussed below are some aspects that have had a profound influence on the development of instrumentation for materials science and where it is expected that further investments will continue to rapidly advance the state of the art.

Experience with state-of-the-art equipment is essential for the training of materials scientists. Major investments in instrumentation were perceived to be necessary by the five industrial members of the Panel to meet the need for well-qualified Ph.D. graduates as well as for new scientific discoveries.

Characterization Techniques

The development of ultra high vacuum technology has also promoted progress in surface science and surface analysis. By being able to keep a surface clean for several hours it is possible to make a thorough characterization of that surface and to do experiments on it. The extension of UHV technology to other characterization techniques, e.g., electron microscopy, is just emerging. The possibility of combining surface analytical tools such as Auger and photoelectron spectroscopy with electron microscopy and energy loss spectroscopy in an UHV environment is an instrumentation opportunity with high leverage now being realized at a few institutions.

Many materials characterization techniques are based on beam spectroscopies. A beam of photons or particles impinges on the sample, and outgoing photons or particles are analyzed with respect to energy, intensity, polarization, and angular distribution. Progress in condensed matter physics in the understanding of beam interactions with matter has promoted the use of these techniques as analytical tools. There are many opportunities for further progress in this field with a strong coupling between the physics and the materials science communities.

The further development of new sources of electromagnetic radiation such as synchrotron radiation sources, high-power lasers, and free-electron lasers provide major opportunities for applications in materials science. The rapid development of surface-sensitive analytical techniques will be further enhanced by access to these sources.

The sensitivity of some of these techniques has been pushed to a spectacular degree. Trace impurities in microsized regions at the level of 1 part in 10^8 can be detected in favorable cases. In sharp contrast to these advances stands the fact that we are still looking for good analytical tools for the analysis of light elements. Surface analysis has advanced to where fractions of a monolayer of elements heavier than helium can be detected. However, the detection of light elements even in the percentage range becomes difficult when they are embedded in the bulk. The situation appears to be particularly cumbersome for insulators and organic-based materials systems.

National Facilities

The generation of some of the beams mentioned previously require large facilities, such as nuclear reactors or accelerators for neutron beams, electron storage rings for synchrotron radiation, and high-voltage facilities for electron microscopy. Continued optimal use of these facilities for materials science depends on continued instrumentation development.

Synchrotron radiation sources provide such capabilities as small angle scattering, crystallography, topography, x-ray microscopes, microprobes, EXAFS, and photoemission spectroscopy. These can be extended with the development of devices such as wigglers and undulators that make it possible to study time-dependent phenomena. Some of the necessary instrumentation developments are made possible by progress in materials fabrication such as artificially structured x-ray monochromators and x-ray transmission optics.

Neutron sources provide unique capabilities for the study of materials. Magnetic scattering, deuterium substitution, inelastic scattering, and small-angle scattering are complemented by similar x-ray methods. Instrumentation advances will provide new methods for high-energy resolution of particular importance in the study of polymers, for example.

It is obvious that major national facilities must serve wide geographical areas. However, even more modest instrumentation facilities like electron microscopy, NMR, surface analysis, and ion beam accelerators have reached such levels in cost and operation sophistication that a certain centralization is necessary. The National Science Foundation's regional instrumentation centers are experiments to try to solve some of these problems, and they should be watched closely to allow an evaluation of the success of this approach. Similarly, the materials research laboratories serve an important function in this respect on a campus basis.

Integrated Materials Preparation and Characterization Systems

The ultimate coupling between materials synthesis and materials characterization steps is achieved by combining the two in the same system, allowing materials characterization to be done simultaneously with the materials growth as is done, for example, in MBE growth. This trend will be further enhanced with the widespread use of microcomputer-controlled equipment and characterization tools. An opportunity exists in the development of appropriate sensors to monitor the process parameters.

Computers

Access to powerful computers will continue to be a prerequisite for further progress in the theoretical understanding of materials. It is interesting to note that today it is possible to compute the ground-state energy of a given structure and to predict the phase diagram of certain multicomponent systems. Access to large computers is also a necessity for the successful modeling of materials systems and processes. The ultimate goal is to obtain such a good model of a material system that an accurate prediction of its properties can be made before it is produced.

As a result of the electronic revolution (partly enabled by the spectacular advances in materials science), we have witnessed a dramatic decrease in computational cost per operation and the development of powerful computers on a chip basis. With access to modern custom design and fabrication tools it is even possible to make "custom" computers in the form of special chips designed for special types of computation, all within reasonable cost limits. A rapid assimilation of this technology applied to materials science represents a unique opportunity.

Computers in the form of microprocessors will play an increasingly important role in the data acquisition and process control of materials fabrication. The rapid growth of computer networks will allow remotely controlled experiments at the large facilities. These networks will play a major role in the dissemination of information as well as in allowing access to powerful information database systems. A challenging task is the training of materials scientists in the use of modern electronics and computers.

Analysis of Hazardous Materials

Although the materials science issues of hazardous materials are not examined in this report, a few comments on their analysis are in order. Safety levels have moved downward as more sensitive analytical tools have become available.

With the advances in sensor technology coupled with microprocessors and data links, it is now possible to monitor very closely potentially hazardous environments and to be able to take appropriate measures to avoid accidents and catastrophes. There remains a strong need for rapid materials analysis and evaluation in awkward field situations, such as a chemical spill on a freeway.

State-of-the-Art Instrumentation

The successful development of instrumentation requires a strong coupling of the application of physical phenomena; the use of modern measurement technology; and the needs of, for example, the materials scientist. An opportunity exists in creating an environment that increases this coupling. To wait until commercial instrumentation becomes available will often lead to a missed opportunity. This coupling usually occurs naturally in an industrial research laboratory environment but to a considerably lesser degree at American universities.

Cost of Instrumentation

With the development of more sophisticated instruments and because of the higher rate of inflation for instruments, the cost of instrumentation has increased sharply. This has posed a serious problem, especially in universities, in keeping up with instrumentation development. A partial solution to this problem is the use of centralized facilities on a regional or campus basis. However, progress in materials science today requires a substantial instrumentation budget for each individual scientist.

CONCLUSION

Opportunities in materials science have been discussed from the viewpoints of the technologies involved, the new materials and processes that can be foreseen, and the challenges to the basic sciences and engineering disciplines that constitute materials science.

Certain themes are pervasive. A better understanding of surfaces, interfaces, and grain boundaries can be achieved, which will lead to more control over synthesis and to new materials with improved properties. This in turn will affect all processes dependent on strategic materials and on the structural, power, and transportation industries. It will lead to new, faster, and cheaper microdevices for the data processing and communications industries. Much of the progress foreseen depends on realizing opportunities for improving instrumentation and its attendant technology.

Concepts and models being pursued in the basic science underpinning of materials science give intellectual stimulation, provide excellent university-based research projects for the training of the graduate students needed by industry, and occasionally provide entirely unforeseen innovations.

RESEARCH BRIEFINGS 1983

for the Office of Science and Technology Policy,
the National Science Foundation,
and Selected Federal Departments and Agencies

Committee on Science, Engineering,
and Public Policy

National Academy of Sciences

National Academy of Engineering

Institute of Medicine

NATIONAL ACADEMY PRESS
Washington, D.C. 1983

National Academy Press **2101 Constitution Avenue, NW** **Washington, DC 20418**

NOTICE: The National Academy of Sciences was established in 1863 by Act of Congress as a private, nonprofit, self-governing membership corporation for the futherance of science and technology for the general welfare. The terms of its charter require the National Academy of Sciences to advise the federal government upon request within its fields of competence. Under this corporate charter, the National Academy of Engineering and the Institute of Medicine were established in 1964 and 1970, respectively.

The Committee on Science, Engineering, and Public Policy is a joint committee of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. It includes members of the councils of all three bodies.

This work was supported by the National Science Foundation under Contract PLN-8218330.

Library of Congress Catalog Card Number 83-63234

International Standard Book Number 0-309-03437-X

Copyright © 1983 by the National Academy of Sciences

No part of this book may be reproduced by any mechanical, photographic, or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted, or otherwise copied for public or private use, without written permission from the publisher, except for the purposes of official use by the United States Government.

Printed in the United States of America

Committee on Science, Engineering, and Public Policy

GEORGE M. LOW, President, Renssalaer Polytechnic Institute (*Chairman*)
LINDA H. AIKEN, Vice President for Research, The Robert Wood Johnson Foundation
JACOB BIGELEISEN, Leading Professor, Department of Chemistry, State University of New York, Stony Brook
FLOYD E. BLOOM, Director, Arthur V. Davis Center for Behavioral Neurobiology, The Salk Institute
EMILIO Q. DADDARIO, Wilkes, Artis, Hedrick and Lane, Attorneys at Law
GARDNER LINDZEY, President and Director, Center for Advanced Study in the Behavioral Sciences
EDWARD A. MASON, Standard Oil Company, Amoco Research Center
JOHN L. McLUCAS, President, World Systems Division, Communications Satellite Corporation
ELIZABETH C. MILLER, WARF Professor of Oncology, McArdle Laboratory for Cancer Research, University of Wisconsin
GILBERT S. OMENN, Dean, School of Public Health and Community Medicine, University of Washington, Seattle
LEON T. SILVER, Division of Geological and Planetary Sciences, California Institute of Technology
HERBERT A. SIMON, Professor of Computer Sciences and Psychology, Department of Psychology, Carnegie-Mellon University
I. M. SINGER, Professor, Mathematics Department, University of California, Berkeley
JOHN B. SLAUGHTER, Chancellor, University of Maryland
F. KARL WILLENBROCK, Cecil H. Green Professor of Engineering, School of Engineering and Applied Science, Southern Methodist University

Ex Officio

FRANK PRESS, President, National Academy of Sciences
ROBERT M. WHITE, President, National Academy of Engineering
FREDERICK C. ROBBINS, President, Institute of Medicine

COSEPUP Staff

ALLAN R. HOFFMAN, Executive Director
BARBARA A. CANDLAND, Administrative Assistant

Preface

In 1982, on an experimental basis, the National Academies of Sciences and Engineering and the Institute of Medicine, through their joint Committee on Science, Engineering, and Public Policy (COSEPUP), developed research briefings on seven fields of science for the White House Office of Science and Technology Policy (OSTP), the National Science Foundation (NSF), and other interested federal departments and agencies. The seven fields were Mathematics, Atmospheric Sciences, Astronomy and Astrophysics, Agricultural Research, Neuroscience, Human Health Effects of Hazardous Chemical Exposures, and Materials Science. Each briefing, developed by a panel of experts, identified research areas within the field that were likely to return the highest scientific dividends as a result of incremental federal investment in fiscal year 1984. Briefing papers were prepared by each panel and published by the National Academy Press in early 1983.

Shortly thereafter, OSTP and NSF requested a second round of research briefings focused on federal investment opportunities in the fiscal year 1985 budget. COSEPUP agreed to this request and with the assistance of units of the National Research Council and through consultations with OSTP and NSF selected five new research briefing topics. These were as follows:

1. *Selected Opportunities in Chemistry*: Research directed toward enhanced understanding and application of three promising areas of chemistry: understanding chemical reactivity, chemical catalysis, and chemistry of life processes.
2. *Cognitive Science and Artificial Intelligence*: Research directed toward understanding how intelligent systems, animate or inanimate, deal with problems of communication, the acquisition of knowledge, and adaptive behavior.
3. *Immunology*: Research directed toward understanding the molecular, cellular, and intracellular processes of the immune system with a view to understanding control of the immune response and functions of the system.
4. *Solid Earth Sciences*: Research directed toward obtaining a new level of understanding of the structure, composition, energetics, and evolutionary history of the earth, with emphasis on the continental lithosphere and its margins.
5. *Computers in Design and Manufacturing*: Research directed toward understanding and

enhancing the application of computers to product design and manufacture.

This document contains the reports developed on each of these topics.

The second round of briefings was patterned after the first. Each one-hour briefing was developed by a panel of about 12 experts who met once for several days to assess the field and identify areas of unusual scientific opportunity. At each panel meeting a rapporteur was present to summarize the discussions and prepare an initial draft of a briefing paper. After further review by the panels, the briefing papers served as the bases of one-hour oral briefings presented to COSEPUP for review in early August 1983. Based on this review, the briefing papers were revised and served as the bases for oral briefings presented to federal officials in September and October 1983. Separate briefings were provided to Dr. George Keyworth, Science Advisor to the President, and selected members of his staff; Dr. Edward Knapp, Director of the National Science Foundation and senior members of the NSF staff; and senior representatives of other interested federal departments and agencies.

Funds to support both rounds of briefings were provided by the National Science Foundation.

George M. Low, *Chairman*
Committee on Science,
Engineering, and Public Policy

Contents

Report of the Research Briefing Panel on the Solid Earth Sciences 1

Report of the Research Briefing Panel on Cognitive Science and Artificial Intelligence 19

Report of the Research Briefing Panel on Immunology 37

Report of the Research Briefing Panel on Computers in Design and Manufacturing 49

Report of the Research Briefing Panel on Selected Opportunities in Chemistry 65

Report of the Research Briefing Panel on the Solid Earth Sciences

Research Briefing Panel on the Solid Earth Sciences

Charles L. Drake (*Co-chairman*), Dartmouth College
Don L. Anderson (*Co-chairman*), California Institute of Technology
William R. Dickinson, University of Arizona
Carl Kisslinger, University of Colorado
John C. Maxwell, University of Texas at Austin
V. Rama Murthy, University of Minnesota
Jack E. Oliver, Cornell University
C. Barry Raleigh, Lamont-Doherty Geological Observatory
Frank M. Richter, University of Chicago
Eugene M. Shoemaker, U.S. Geological Survey
Edward M. Stolper, California Institute of Technology
Peter J. Wyllie, California Institute of Technology

Staff

Pembroke J. Hart, *Executive Secretary*, Geophysics Research Forum
Barbara Valentino, *Staff Associate*, Geophysics Research Forum
Patricia Miron, *Secretary*
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

Report of the Research Briefing Panel on the Solid Earth Sciences

SUMMARY

The Panel on Solid Earth Sciences has selected five broad targets of opportunity that have the capacity to provide rapid advances in the near future and that will contribute most to our understanding of the Earth's interior and history. The common themes are the structure, composition and evolution of the continental lithosphere, and the dynamics of tectonic processes. This report emphasizes areas that lie on the frontiers of earth sciences—in some cases on the frontiers of chemistry and materials sciences as well—and describes some of the conceptual and technical advances that make it possible to explore more fully the third and fourth dimensions, depth and time. With additional resources, available skills and techniques can be applied immediately to some of these basic problems. Some projects, such as deep continental drilling, determination of the continental geoid, and crustal seismic reflection, can proceed immediately if resources are made available. Others, such as large seismic arrays, expanded isotopic exploration of the crust and mantle, monitoring of crustal motions, and the chemistry and physics of geological materials, require major investments in modern facilities.

With modern facilities and instrumentation the basic research opportunities of these projects will provide excellent training for the next generation of scientific leadership in the earth sciences as well as the advanced skills and trained personnel sought by a wide variety of industries, especially those seeking natural material and energy resources.

INTRODUCTION

Twenty years ago the hypothesis of seafloor spreading was proposed. It led to general acceptance of the concept of a dynamic Earth and to the plate tectonics model as an explanation for the principal tectonic features. In this model the rigid outer shell of the Earth, the lithosphere, is broken into a limited number of large plates moving relative to one another. Where the plates separate at the ocean ridges new lithosphere is created, which thickens with age and cooling. Where the plates converge—the island arcs and young mountain systems—old ocean lithosphere returns to the deep mantle. Where the plates slide against each other—as in California—

major fault systems develop. Accompanying the interactions of the plate margins are earthquakes, volcanoes, and the concentration of mineral resources.

The importance of this model to the geological sciences can hardly be overstressed. It provided earth scientists for the first time with a working model of the Earth as a whole, a unifying concept of global structure and composition, a fresh context in which to view earth history, and a framework in which to set detailed local investigations.

In succeeding years, development of the model has led to very successful explanations for the development of the oceanic lithosphere and of the major topographic features of the oceans. Through comparative studies of the moon and the terrestrial planets, we have generated a picture of the early history of the Earth and why it appears to be the only one of these bodies on which plate tectonics processes have been active. We are less certain of the nature of the driving forces that move the plates, and we are only beginning to understand the manner in which the continents have been generated and assembled over the last 4 billion years.

In the last decade we have come to recognize that the continents are made up of a large number of microplates assembled over long periods of time. We have found that hydrothermal circulation penetrates deep into the crust and that on ocean ridges the return circulation to the surface may yield metallic sulphide deposits in remarkably short periods of time. We have discovered evidence of major singular events that may have affected the geological record and life on a global scale. We have become able to evaluate the heterogeneity of the mantle, both in physical properties and in chemistry, which has implications with regard to the driving mechanism of plate tectonics and to the formation of continents. In the framework of plate tectonics we have been able to construct generic models to assist us in our search for resources and for a better understanding of the causes of natural hazards.

Technological advances have accompanied these conceptual advances. We can now measure the properties of materials at pressures comparable to those at the center of the Earth. We can perform precise geochemical and isotopic analyses of very small samples of earth materials at levels of 10^{-9} to 10^{-10} grams. Techniques are becoming available that will allow us to monitor crustal motions and positions at the centimeter level of precision. Space techniques have been developed that have permitted us to map the gravity field over the oceans in surprising detail, revealing the fine structure of the sea floor. We can apply tomographic techniques to map the three-dimensional structure of the Earth's interior.

The diversity of approaches within the academic earth sciences is generally well served by the research grant programs of the National Science Foundation (NSF) and, to a lesser extent, of other federal agencies. The Panel strongly supports these programs. In addition, there now exist research opportunities, often large in scale, that are especially timely because technology has become available or because significant progress demands a new approach. Such programs, e.g., scientific drilling or establishment of seismic arrays, are costly and may require considerable investigator collaboration and coordination, but they offer the promise of major increases in our knowledge of the Earth, of how it works and has worked, and of its internal resources.

The scientific problems upon which the recommendations of the Briefing Panel are focused are as follows: (1) *Structure and Composition of the Lithosphere*; (2) *Dynamics of Tectonic Processes*; and (3) *Evolution of the Continental Lithosphere*.

The Panel has based its recommendations on relevant reports by committees of the NAS/NRC, especially the report *Opportunities for Research in the Geological Sciences* prepared at the request of the Director of the Division of Earth Sciences of the National Science Foundation (NSF).

Geoscience research activities of industry and of federal agencies are often closely intertwined with those of academia, although the goals may differ. Industry has developed technologies and data that are useful in enhancing our basic understanding of the Earth, and it has also applied ideas and techniques developed in academia to its own problems. Many university research activities supported by NSF are dependent on facilities supported or operated by other federal agencies; without these facilities the research cannot be done. At the same time, without support for this research, full advantage cannot be taken of the facilities and graduate students cannot be trained.

The Briefing Panel has identified five research areas in which significant dividends can be expected as a result of incremental federal investment in FY 1985. These five research areas are:

1. Seismic Investigations of the Continental Crust.
2. Continental Scientific Drilling.
3. Physics and Chemistry of Geological Materials.
4. Global Digital Seismic Array.
5. Satellite Geodesy.

ANALYSES OF RECOMMENDATIONS

SEISMIC INVESTIGATIONS OF THE CONTINENTAL CRUST

Seismic investigations include both reflection and refraction seismic surveys, which complement each other. Reflection probing of the continental crust has opened the third dimension to detailed view and has revised our concept of the deep structure of the continental crust and upper mantle. One technique is a modification of petroleum exploration methods. Because of the large service industry involved in the search for petroleum, availability of funds is essentially the only barrier to expansion of this type of exploration. At present only one field crew is being employed full time surveying for scientific purposes in the United States. An augmented effort would provide an opportunity to investigate some of the many attractive target areas that have already been identified as well as to experiment with new ways to apply the technique to obtain greater detail and increased penetration.

Seismic refraction and imaging investigations can use natural (earthquake) or artificial (explosive) sources to probe the continental lithosphere. In this case the research is presently equipment limited. Closely spaced arrays of up to 1000 seismic instruments are required to sample the lithosphere on a scale comparable to its geological heterogeneity. Operation of such an array requires considerable coordination and planning, and these have already begun. Capital investment in portable digital instruments should follow at a rate of 100 instruments per year, and when sufficient instruments have become available a large-scale experiment should begin. Some experiments using new three-dimensional imaging techniques will site the array for several months over particularly interesting areas for determination of deep structure.

The transition from continent to ocean is the fundamental discontinuity in the crust. It takes two major forms: passive margins, which are intraplate, seismically quiet transitions from continental to oceanic crust, and convergent margins, with earthquakes and volcanics, where oceanic lithosphere interacts with that of the continent. The U.S. Geological Survey (USGS), industry, and academic institutions have probed the upper few kilometers of both types of margins, but the deeper structure and the tectonic processes remain a mystery. Innovative applications of existing technology promise to delineate the details of the deeper structure and to allow us to reconstruct the structural, tectonic, and thermal history of these areas.

The most incisive experiments are provided by large seismic (reflection and refraction) arrays. These will require time on ships equipped with seismic reflection instrumen-

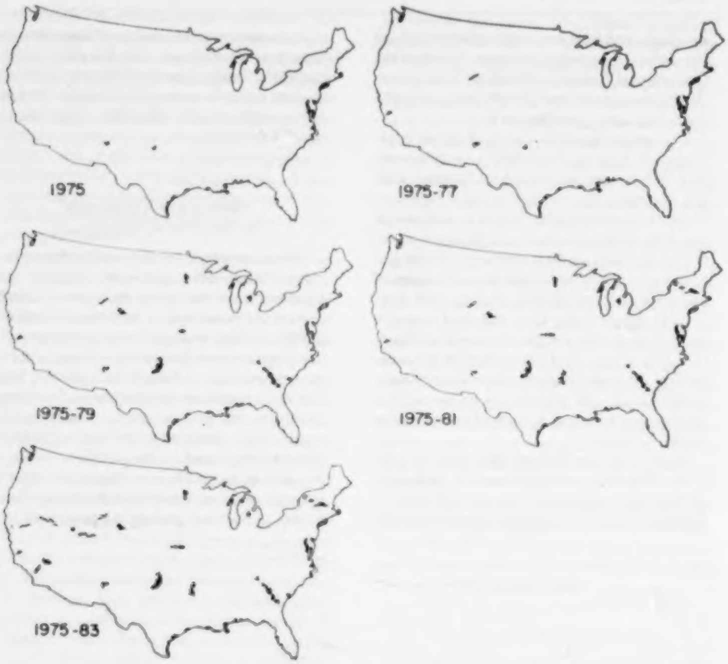


FIGURE 1 COCORP (Consortium for Continental Reflection Profiling) survey lines now total approximately 5200 kilometers.

tation comparable with that available to industry and arrays of ocean bottom seismometers similar to the continental array discussed earlier. On passive margins the continental crust extends as much as 200 kilometers offshore and often contains large sedimentary basins with resource potential. Seismic reflection surveying at sea will provide answers to critically important questions about the origin and evolution of these basins.

CONTINENTAL SCIENTIFIC DRILLING

Drilling in the oceans has revolutionized our ideas about the history and development of these areas; continental scientific drilling promises to be equally exciting.

In the report *Continental Scientific Drilling Program* (NAS/NRC, 1979), four principal problem areas for drilling are identified: thermal regimes; ore-forming processes; understanding earthquakes; and the composition, structure, and time and space relationships of the rocks of the continental crust. Panels of the Continental Scientific Drilling Committee (NAS) have identified priority targets in the these four areas (see Figure 2).

The report *Opportunities for Research in the Geological Sciences* (NAS/NRC) recommends that NSF-funded dedicated continental drilling be initiated in support of numerous research opportunities. This is in addition to scientific drilling carried out by other agencies in support of their missions. Drilling is costly and must be based on well-founded plans for research before, during, and after drilling operations; thus the first requirement is for a national planning and preparation effort for holes dedicated to science. Several regional consortia have already been established, and efforts are under way to establish a national group of which these would be a part.

Many holes are drilled each year for purposes other than scientific research, and some of these are so planned and located that significant information could be gained from add-on experiments. Several such experiments have been carried out, but the absence of a dedicated budget for such experiments has made it difficult to respond to many of these opportunities in the short time frame that is usually available.

PHYSICS AND CHEMISTRY OF GEOLOGICAL MATERIALS

Measurement of the physical and chemical properties of rocks, minerals, magmas, and aqueous solutions under the extreme conditions of temperature and pressure found in the Earth's interior is required to understand the composition and flow in the mantle, the origin and evolution of magmas, and the evolution and circulation of hydrothermal systems. Similarly, the precise measurement of isotopic ratios of many elements, with techniques developed in the past few years, is needed to trace the secular evolution of igneous rocks and to establish the time constants of various geological processes. An

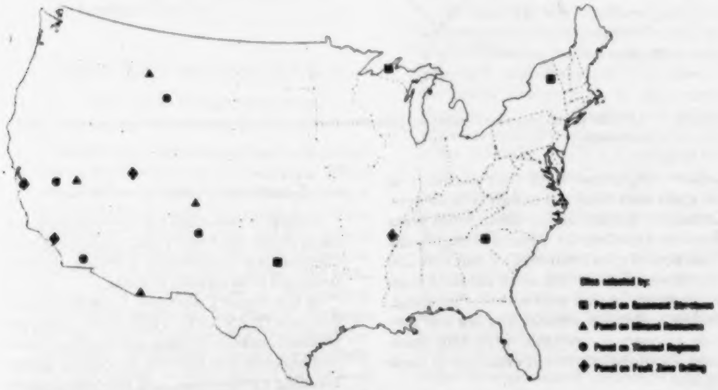

Sites selected by:
■ Panel on Basement Structures
▲ Panel on Mineral Resources
● Panel on Thermal Regimes
◆ Panel on Fault Zone Drilling

FIGURE 2 Proposed drill sites, continental scientific drilling committee.

extremely powerful set of techniques for identifying the major chemical reservoirs in the crust and upper mantle tracing materials through transfer processes and for estimating the age of formation of their derived products has been developed. These techniques utilize isotope chemistry and a variety of mass spectrometric instrumentation. A systematic survey of major global rock isotope and geochemistry systems would provide fundamental solutions to many major geological problems.

Progress in these areas depends not only on access to modern technologies and instrument development in the laboratories of chemistry, physics, and material sciences but also on the continuing development of these techniques to make laboratory measurements under the conditions of temperature, pressure, and natural contamination that characterize many geological environments. The high cost of instrumentation in this area and the limited funding available have inhibited the exploitation of rich opportunities.

This country has been a leader in developing the concept of plate tectonics, in applying innovative tools and techniques to studying the Earth and its neighboring planetary bodies, and in developing and manufacturing new instruments. In recent years this leadership posture has been eroded to some degree. A vigorous research program supported by the atomic energy and space efforts equipped many academic laboratories, but no comparable effort in instrumentation has been made for more than a decade. U.S. facilities now show significant obsolescence. If advantage is to be taken of the opportunities that are available in geochemical, isotopic, and high-pressure studies, a major commitment must be made to the development, purchase, and maintenance of new instruments and to the modernization of laboratories.

Global Digital Seismic Array

The seismograph array is the antenna by which the geophysicist receives signals propagated by the Earth. The global system of analog instruments has been the main source of data for seismological studies. A small number of digital installations are now in use, supported by DOD, USGS, and NSF; they have confirmed the enormous power of digital data for solving important problems. The goal is to determine the variation of velocity, anelasticity, anisotropy, and stress state through the Earth and, eventually, to map composition and convection.

This goal can be accomplished through a calibrated set of broad-band instruments emplaced around the surface of the globe. An essential element is transmission (near real-time) of the data from about 100 stations by satellite telemetry to a properly equipped data center. All of the technology for achieving this network now exists, but significant support will be required to deploy the system, establish the central receiving and distribution center, upgrade a complementary analog facility, and fund the necessary supporting research.

The Panel recommends that the National Science Foundation act as overall coordinator and lead agency for funding such an array and that the operation be planned and overseen by a university consortium.

Satellite Geodesy

The new Global Positioning Satellite (GPS) system promises resolution of position to the centimeter level, giving us the capability of measuring actual plate motion and deformation rather than the results of these integrated over time. Use of GPS will revolutionize both conventional and tectonic, or time-dependent, geodesy.

To take advantage of this new tool, an investment will be required for construction of water-vapor radiometers and to purchase an initial set of GPS receivers for testing the accuracy against existing optical measuring instruments. In subsequent years support will be needed for the acquisition of additional instruments, for field observations, and for the operation of a satellite orbital determination network.

Satellite altimetry (measuring the elevation of the sea surface) recently has been used to construct maps of gravity anomalies in the oceans. These anomalies demonstrate, for example, that fracture zones are much more profound features than had previously been recognized, and they provide direct evidence for shallow convection in the mantle. These measurements will be continued by TOPEX, which is designed principally for oceanographic purposes. The resolution of these measurements can be improved by the systematic measurement of the gravity field from space, as proposed in the Geopotential Research Mission (GRM) of NASA. In addition, a mission such as this can extend high-resolution gravity (and magnetic) measurements to the continents. The value of this mission in providing global gravity coverage and better orbital data for TOPEX and other earlier missions make it an attractive opportunity for an early start.

THE SCIENTIFIC OPPORTUNITIES THAT CAN BE EXPLOITED BY THESE RECOMMENDATIONS

STRUCTURE AND COMPOSITION OF THE CONTINENTAL LITHOSPHERE

In the last few hundred years the two-dimensional surface of the Earth has been explored both geographically and geologically. In this century the third dimension, depth, has been probed intensively in selected areas, particularly the oceans and sedimentary basins, but with limited resolution. Achievements with major impacts on science and society such as development of the concept of plate tectonics and the discovery of major deposits of mineral and energy resources followed this exploration. In this historical context, the next great frontier to be explored clearly seems to be the great body of rocks that form the continental crust and mantle.

Although many modern geological, geophysical, and geochemical studies are compatible with a comparatively simple layered crust, there are increasing indications of complex and discontinuous multilayers, some complexly deformed internally. Exposures of deep crust and samples of deep crust brought up by volcanoes indicate that a variety of sedimentary and volcanic rocks of surficial origin may make up a significant part of the lower as well as the upper crust.

In recent years it has become possible to apply the high-resolution geophysical techniques developed for petroleum exploration to the study of the entire thickness of the continental crust. Although, geographically, only a fraction of the crust has so far been examined, the results reveal a crust of far greater complexity than previously envisioned. In orogenic belts, thin slices of older rock have been thrust for hundreds of kilometers over younger sedimentary rocks of preexisting continental margins. This is an effect of the cycles of opening and closing of ocean basins that seems secondary in terms of basic plate movements but is of primary importance to the development of the continental crust. In areas of active rifting, such as the Basin and Range Province, steeply dipping faults forming and bounding the ranges are found to flatten and to merge at depths with shallow detachment faults that stretch for a hundred or more kilometers beneath the extensional terrain. Such faults may be reactivated zones of weakness from still earlier episodes of major compression and thrust faulting. What is perceived as the fundamental boundary between crust and mantle, the Mohorovicic discontinuity, is clearly far more complex and variable both in vertical and horizontal directions than was commonly thought, and it may vary widely in nature from one locality to another. Faults that may be conduits for mineral-bearing fluids have been traced from surface features, including areas of mineral concentration, to great depths. Molten bodies of rock (magmas) have been mapped at crustal depths where solid rock normally prevails. Deeply buried sediments that are worthy of exploration for petroleum have been found.

The continental crust does not disappear at the shoreline; it extends beneath the continental margins. The margins of the continents vary according to the nature of the processes that have shaped them. Rifting apart of the Atlantic Ocean has left behind a continental shelf in the United States where the evaluation for the potential of oil is dependent on better exploration by refraction and reflection seismic methods. The western margin of Central and North America contains the entire panoply of active plate tectonic processes; rifting in the Gulf of California, subduction of the Pacific plate beneath Central America and the Aleutians, and transcurrent motion from Mexico to northern California. Modern marine seismic methods have advanced rapidly, and these regions are now ripe for intensive exploration.

These new and striking observations of the crust revealed by seismic reflection profiling clearly indicate that a greatly enhanced understanding of the continents will be derived from further observations. Complementary information is derived from wide-angle reflection measurements, refraction measurements on the continents, and the application of newer imaging techniques. To provide measurements with adequate resolution, it is necessary to deploy portable instruments in very large, dense arrays, with up to 1000 instruments, much larger arrays than are currently available or that have been used in past investigations of the lithosphere. These arrays may be employed in refraction profiling using explosive sources, or they can be deployed to monitor earthquake activity over periods ranging from days to months. Such monitoring can provide locations of hypocenters so precise as to permit correlation with specific geological structures at a new level of precision—a level that would significantly advance the understanding of earthquakes and the tectonic processes that cause them. Distant seismic events recorded on these arrays provide the data for deeper imaging.

The ultimate exploration tool, as well recognized by the mineral and petroleum industries, is the drill. Deep-sea drilling has been invaluable in providing verification of geophysical interpretations and has provided detailed information on the structure, composition, and history of the ocean basins. A number of countries have established continental scientific drilling programs, including Canada, France, Belgium, Japan, West Germany, the United Kingdom, Czechoslovakia, and the USSR. The program in the USSR includes a scientific hole in the Kola Peninsula that is now at a depth of 13.5 kilometers (44,000 feet); drilling is continuing with a target of at least 15 kilometers. This is the deepest hole in the world (previous record was 33,000 feet in Oklahoma) and is located entirely in ancient Precambrian rocks. Most of the holes drilled in this country are drilled for water, oil, or minerals. Many are in anomalous areas, where resources are likely to be found; consequently, the basement rocks of the continental crust are unknown in large areas of the country. While some knowledge can be gained by attaching add-on experiments to the holes drilled for other purposes, it is becoming apparent that dedicated holes will be needed to fully understand the structure, composition, and age of the continental crust.

The continental crust and deeper parts of the lithosphere contain the record of about 4 billion years of earth history. The continental crust is also the repository of recoverable resources of minerals and fuels. Many segments of the present continental crust date from a time far back in geological history. There are good physical and chemical reasons to suppose that tectonic regimes and geodynamic mechanisms have changed with time. Important questions arise such as when and how rapidly the continental crust was generated, whether the process was episodic or continuous, and the rate of generation as a function of time.

Studies based on modern isotopic techniques suggest that we are at the threshold of being able to answer some of these questions. It now appears that the continental lithosphere is in some sense complementary to a residual mantle region of global scale. The rate and manner of extraction of continental crustal material from mantle sources can be understood by extensive isotopic and chemical studies of rocks from the upper and lower continental crust. These measurements provide the ages of and information on the nature and history of source materials from which the continental crust has been derived. Present isotopic data suggest that the formation of the continental crust is episodic and that very large volumes are formed in relatively short times.

Finally, in recent years it has been found that the differences between continents and oceans extend to considerable depths (150 kilometers or more) and that the deep structure of the mid-ocean ridges may extend to 400 kilometers. It is not clear whether these differences are statically or dynamically produced, but it is clear that understanding the nature of the continents will require exploration to great depths beneath them. Fortunately, new seismic techniques, similar to medical tomography, can provide this kind of information, both on a regional and a global scale.

DYNAMICS OF TECTONIC PROCESSES

Contemporary tectonic theories provide a framework for the synthesis of a wide variety of geological, geophysical, and geochemical data. However, these models are primarily descriptions of operative processes and are only slowly developing into rigorous and quantitative formulations. Increased amounts of new and high-quality data; advances in continuum mechanics, materials science, and other relevant physics; and availability of adequate computers make the present time right for rapid progress toward the development of more effective models of structural processes.

Research in tectonics is carried out in the field, where nature is observed; in the laboratory, where the myriad parameters in the real world can be controlled; and through theoretical studies. Laboratory studies of rock deformation tell us about processes in rocks and the changes that result from high deviatoric stress. The changes caused by stress provide a basis for interpreting geodetic and geophysical data in tectonically active areas. Seismic techniques are being applied to the determination of magnitude and orientation of stress in the mantle. Studies of petrofabrics in the field and in the laboratory provide clues to the temperature, stress, and flow conditions that prevailed during the formation of the rock. Thus we can study past as well as contemporary dynamics of the lithosphere and the underlying mantle.

The data critical to establishing the kinematics of plate motion were based on paleomagnetic studies and marine geophysical measurements. Now, with the remarkable accuracy of new Global Positioning Satellite receivers, we can foresee that within a few years plate velocity vectors can be determined and accelerations in plate motion can be measured. The interaction of the plates in great earthquakes can be observed directly by measurement of earthquake displacements and post-seismic dissipation of strain. This new, relatively inexpensive technology provides one of the most exciting opportunities in the earth sciences today.

Knowledge of the stresses in the lithosphere is fundamental to understanding the physics of tectonic processes. Because it is a tensor quantity describing forces internal to a mass of material, stress is difficult to observe directly. Overcoring and hydrofracturing techniques have been used for in situ measurements of stress in rocks, but only the latter is useful at significant depths. The cost of hydrofracturing rises steeply with depth because of the need to drill a hole; therefore, stress measurements should be included in the observation program for all deep holes drilled for scientific purposes. New holgraphic techniques are being developed for down-hole applications.

Stress measurements at isolated points are the beginning step in analysis, but it is the spatial variation in stress that drives dynamic processes. Because earthquakes are the result of stresses at the source, both the radiated waves and the permanent deformation they produce contain important information on stress. Seismic source theory and wave propagation theory are sufficiently developed that the dynamic and geometric properties of the sources can be derived routinely from analysis of sets of seismograms. This is done in the time domain by matching the calculated and observed waves and in the frequency domain by interpreting the spectral characteristics. The greatest need is for high-quality digital data from earthquakes with a wide spread of geographical distribution and magnitudes. The technology exists; it is a matter of deploying the instruments and creating the data management facilities to serve the experimenters.

Digital data are also required for the analysis of elastic wave velocity anisotropy. Interpretation of anisotropy data is more ambiguous than waveform analysis because it can result from both deviatoric stresses and crystal alignment. Crystal alignment in rocks of the mantle is itself the result of flow; thus, it can shed light on internal dynamic processes. These anisotropy studies have the advantage that they can be used to probe any region through which the waves travel, not just tectonically active areas. Detection of anisotropy from wave polarization anomalies and surface wave dispersion is being recognized as a powerful tool for studying Earth dynamics; the availability of large quantities of broadband digital data will enhance development of this approach. Similarly, the attenuation of seismic waves is controlled by temperature and dislocation structure and, thus, stress.

Finally, the response of rocks to stresses is governed by their rheological properties. More complete knowledge of the elastic and inelastic mechanical and thermodynamic properties of the materials in the crust and mantle is needed as a basis for modeling processes ranging from metamorphism to basin formation and mountain building. New laboratories need to be built to measure the elastic and anelastic properties of minerals at extreme conditions and to synthesize materials at high pressure and temperature. These are frontier areas of materials science.

Dynamics and Structure of the Earth's Interior

The Earth's mantle and core are the major components of its interior. From geophysical studies, experimental investigations reaching the temperature and pressure conditions of the core-mantle boundary, geochemical and isotopic analysis of mantle-derived rocks, and modeling, we have obtained a rough understanding of the composition and physical state of the mantle and the core.

Plate tectonics and related phenomena result from flow in the mantle. This flow is also the principal process by which heat escapes from the Earth's interior. Beneath a rigid outer shell—the lithosphere—the Earth's mantle, which responds like a solid to short-term driving forces, behaves as a fluid over geological time. Plate motions are the result of convection in the mantle, but the pattern of convection is controversial; some advocate whole-mantle convection, others suggest separate systems in the upper and lower mantle. It is not known whether mantle convection occurs on a scale that is small compared to the large-scale flow associated with plate motions and whether the time dependence associated with the plate configurations is the only time-dependent aspect of plate tectonics.

It has been difficult to place observational constraints on the flow in the Earth's interior. However, techniques are now available for mapping of the Earth's interior in three dimensions (elastic wave tomography). One approach involves the systematic inversion of large amounts of travel-time data of compressional and shear waves that have traveled through the Earth; the other involves modeling based on free oscillations and surface waves having periods sufficiently long to be affected by the deep interior. Both approaches require large computing facilities and a global array of broad-band digital seismometers. Higher-resolution studies require large transportable arrays.

Recent advances in theory and increases in computational capacity are available to help solve the fundamental flow problems. Most models of mantle flow have been based on a single layer behaving as a Newtonian fluid. Mantle silicates have strong temperature- and stress-dependent rheologies and the mantle is surely inhomogeneous in viscosity and probably in chemistry. Melting that occurs in the course of convection affects the buoyancy. It is now possible to compute flow in a layered system with these complications taken into account; preliminary results are providing insight into the dynamics of the interior. When continents are present on the surface it may be that steady-state never prevails. The required three-dimensional finite element calculations exceed the capacity of most university computers and computer budgets, but they must be made if the understanding of mantle dynamics is to proceed beyond the present qualitative stage.

In addition to seismological data, data on the global geoid, gravity and topography, crustal and sedimentary rock thickness, and heat flow are also pertinent to investigations of mantle flow, rheology, and evolution. Of special value among newly available data are satellite-borne radar measurements of the shape of the sea surface (geoid). There are satellite techniques for extending this high-resolution geoid to the continents. Systematic treatment of these data will make it possible to place bounds on the locations of density and velocity anomalies in the mantle. Explanation of the geoid is one of the most challenging problems of modern earth science. If we understood the geoid, we should know whether the rheology of the mantle is strongly temperature and stress dependent, whether the convection is layered or whole mantle, and whether convection is statistically steady. It has been suggested recently that the geoid has a memory—that its present shape reflects conditions in the distant past. If this is true, then we have the exciting prospect of being able to identify past configurations of the continents, ridges, and subduction zones.

Laboratory measurements of the rheology of melts and silicate minerals at high temperatures and controlled dislocation densities are an essential part of these studies. Experimental geophysics has gone through a decline in recent years, in part because of obsolescence of equipment. New laboratories, filling the gap between rock mechanics and ultrasonics, are required to provide the basis for interpreting seismic attenuation and anisotropy observations in terms of rheology, stress, and flow.

A major advance in our understanding of the mantle has been the discovery that the mantle is chemically heterogeneous on a variety of scales. Evidence of this is provided by the precise analysis of Pb, Sr, Nd, and He isotopes of young volcanic rocks from the oceans. These data show that the mantle sources that yield mid-ocean ridge basalts (MORB) have been isolated and depleted, relative to other sources of basaltic rocks, in their radioactive and other trace elements for a period on the order of 1.5 to 2.0 billion years. Furthermore, the mantle sources that produce MORB appear to be a distinct global entity: relatively homogeneous chemically but distinct from all other mantle sources that produce basalts. For example, basalts from major volcanic centers on oceanic islands and from continental interiors cover a broad geochemical and isotopic range. The sources of these so-called hotspot magmas are enigmatic. Investigations in this area have been limited. Few laboratories are working on these fundamental problems because of the expense of establishing and maintaining modern isotopic facilities and their supporting high-precision separation and analytic equipment.

Modern labs have recently been set up in England, France, West Germany, Australia, and Japan, often by U.S.-trained investigators.

Magma Genesis and Fluid Flow

Igneous activity—melting in the Earth's interior, migration of magmas toward the Earth's surface, eruption on the surface or emplacement at depth—is the major process of crustal formation and one of the major mechanisms of heat and mass transfer in the outer regions of the Earth. It is the principal mechanism by which volatiles are transported to the surface to form the atmosphere and hydrosphere; cooling of magma bodies sets up hydrothermal systems that produce some of our most vital mineral deposits. Volcanic eruption of magmas may present a significant hazard, but the study of the solidified products of magmas provides clues about conditions in the Earth's interior and their evolution. Generation of magmas not only results from convective upwelling but also, through phase transformations, may influence the density structure of the mantle and provide a positive feedback on the convective process itself.

Investigation of magmas requires a strong combination of field and laboratory studies, but the most significant recent advances have come from development and application of tools from physics, chemistry, and material science to the problems. We can now, in the diamond cell, measure the properties of minerals and fluids under static pressures up to 1.5 megabars and from cryogenic temperatures to 3000° C; we can study the speciation and kinetics of aqueous solutions similar in composition to ore-forming hydrothermal solutions; we can study crystalline and amorphous materials from new perspectives using newly developed or accessible spectroscopic techniques and surface techniques from the materials sciences. These developments suggest that future advances will be led by the new types of and higher-quality data made available through application of innovative analytical and experimental techniques.

Theoretical studies of fluid flow have also played an important role in modifying current thinking about processes in the lithosphere. Fluid dynamic modeling of magma migration in the mantle and crust and of circulation of aqueous fluids in the crust, including the progressive chemical interactions taking place during fluid flow and their feedback on the flow processes, are in their infancy. Preliminary results suggest potential breakthroughs in understanding the factors responsible for mineral-deposit formation and many of the observed chemical and eruptive characteristics of magmas. Such models are critically constrained by geochemical and isotopic measurements on "fossilized" circulation systems and magmas and by measurements of permeability of rocks and the physical and chemical properties of rocks, magmas, and hydrothermal fluids. The systematic application of isotopic techniques to the study of natural stable and unstable isotopic systems is critical. Field investigations, including deep sampling by drilling, and laboratory studies of rocks and model systems are needed to provide realistic constraints and inputs into such studies.

High-Pressure Geophysics

Our understanding of the constitution of the Earth's upper mantle is based primarily on interpretations of seismic data and on the study of igneous and metamorphic rocks believed to have been brought up from relatively great depths by volcanic activities. At deeper levels (greater than 200 kilometers), where our only probes are seismic waves and interpretations of the geoid, realistic interpretations in terms of actual mineralogy, composition, temperature, and rheology require measurements of the physical properties of individual minerals and aggregates at temperatures and pressures relevant to the mantle. Elastic constants and their pressure and temperature derivatives, rheological properties, and phase relations are vital to these considerations. Facilities for synthesis of large volumes of high-pressure phases in order to determine thermodynamic and elastic properties are available (principally outside the United States), and new opportunities are now available for studying the physical and chemical properties in the diamond cell under extreme conditions of temperature and pressure. Such measurements include scattering and reemission (e.g., Raman, fluorescence), diffraction (e.g., X-ray, Brillouin), and absorption spectroscopy (e.g., Mossbauer effect, visible, and infrared). Such measurements are essential not only for interpretation of seismic and gravity data but also as input into realistic models of convection in the Earth's interior.

Evolution of the Continental Lithosphere

Our concepts of the manner of evolution of the Earth and its life have been heavily based on the gradualist ideas of Lyell and Darwin. Recently this manner of biological evolution has been questioned, and it has been suggested that it is not gradual, but punctuated; singular events may play a major role in determining the evolution of organisms. When we consider the evolution of continents we find a similar situation. One can explain the shaping of the Earth's surface as the result of slow processes; mountains appear to be built in geologically short intervals of time separated by long periods of relative quiescence. The average integrated rate of plate movements, measured paleomagnetically, is a few centimeters per year, but current activity along the plate margins is episodic, with earthquakes followed by quiescence; periods of intense volcanism followed by periods of little activity. Singular events, such as the great explosions of Yellowstone, Long Valley, or the Valles calderas, located within the continental mass, are anomalous in terms of a simple gradualistic model.

The continents, with their great age and complexity, carry within them not only most of the Earth's exploitable energy and mineral resources but also clues to the Earth's long history. In contrast, although the rocks of the ocean's floor are relatively young (less than 200 million years), they provide us with the keys to plate tectonics. The great ocean rifts are the primary distillery for separation of the lighter fractions of the mantle into the ocean crust.

The ancient cores of the continents are fringed by younger belts that have been added in the last billion years. In some zones of collision, accretion to the continent of shingles of overthrusted oceanic crust and sediment has contributed to continental growth, while the denser mantle has descended below the continental lithosphere. Study of the accretionary process at the margins of the continents needs to be augmented by extensive examination through geological and geophysical methods and, ultimately, by drilling.

The steady accretion process is punctuated by collisions of continents and by episodes of rifting of the continental lithosphere to create new continental margins and new ocean basins. Sedimentary basins, often containing hydrocarbons, form over these rifts, and the sediments contain the evidence of the rifting process and subsequent subsidence. Rifting has been pervasive during the last billion years, but many of the rifts cease their activity and do not become continental margins. The eastern continental margin is an outstanding example of a successful rift. Its shallower portions have been studied extensively for scientific and petroleum exploration purposes, but understanding of its deeper regions and its full history await clarification by advanced seismic techniques and drilling.

The continental blocks now appear to have been assembled through the course of time from small fragments, some broken off other continents. Examples of this phenomenon have been best identified both on the eastern and the western parts of the country but are evident elsewhere. A broad spectrum of techniques can be brought to bear on identification of these fragments: paleomagnetism;

geochemistry and isotope geology; field mapping; structural and stratigraphic studies; and, where the rocks are buried, seismic measurements and drilling.

Prior to one billion years ago, the evidence for modern plate tectonics processes becomes more obscure. This is not surprising, as the record itself becomes more limited. Processes other than those associated with plate tectonics as we understand them may have contributed to the formation of continental blocks. At the present, two opposing hypotheses can be entertained: that the processes of plate tectonics, operating over a long period of time, have caused the gradual emergence of continental crust from the mantle, and that most of the volume of continental crust was generated early in the Earth's history by processes lacking modern analogs.

To answer this fundamental question we will have to draw on all of our tools—geological, geophysical, and geochemical. Within the major parts of the crust, continent and ocean, and the mantle and core, large-scale geochemical heterogeneities or reservoirs have appeared since the formation of the Earth. The number, size, spatial arrangement, and evolution of reservoirs capable of yielding distinctly different magmas constitute a matter of vigorous debate. In addition, the nature and magnitude of fluxes between reservoirs are poorly known but are essential to understanding the chemical and differentiation history of the Earth. For example, some oceanic basalts have been interpreted as retaining geochemical signatures of recycled ancient ocean crust, raising the question of whether differentiation of the mantle is proceeding irreversibly or whether it is approaching a steady state with recycling.

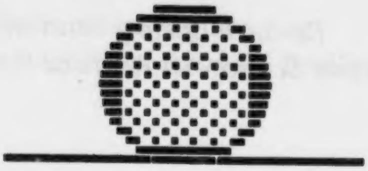
Characterization of the reservoirs can be greatly advanced by the simultaneous measurement of the Nd/Sm and Rb/Sr isotopic systems together with measurements on the same sample of uranium, thorium, lead, and oxygen isotopes and trace-element abundances. A variety of other isotope systems is now being exploited. The measurement of rare gases, particularly ^{3}He/^{4}He, has shown that mantle reservoirs may not be totally outgassed; components that have not previously been on the Earth's surface are still being supplied from the interior. The distribution and evolution of geochemical reservoirs are intimately related to the dynamics of the plate-mantle system and to the thermal and tectonic evolution of the Earth and its continents. The trace-element and isotopic chemistry of crustal and mantle magmas and rocks are essential petrological data, yet only a few laboratories in this country are equipped to make these measurements.

CONCLUSION

Early investigation of the Earth required keen observational powers and a knowledge of the composition of rocks and minerals, of the distribution of organisms through time, and of fundamental geological processes. Through these investigations we gained understanding of the near-surface geology and we were able to exploit this knowledge to obtain resources from the Earth. Similarly, we were able to determine the gross features of the Earth's interior, first through geophysical observations and, more recently, through geochemical and geophysical experimentation and isotopic studies.

The solid Earth interacts with all of the other domains of the geosphere: oceans, atmosphere, solar terrestrial. It is the sole source of information regarding the history of actions and reactions of the geosphere and biosphere. Improved knowledge of the solid Earth and the processes that operate within it are significant not only to basic science but also to society. We live on the solid Earth; we extract our resources from it; it produces the magnetic field through internal motions; from it came the gases of the atmosphere, the waters of the sea and the salts contained within them, and the nutrients that feed its organisms. Terra firma usually describes it well, but internal and external forces produce changes with significant impact on society.

Our tools have become sharper in recent years, and advances in theory and technology provide us with new opportunities to explore in greater detail the nature of and processes in the interior of the Earth as well as the development of the Earth through time. The availability of modern instrumentation and techniques will contribute to the training of the next generation of scientific leadership in the solid earth sciences and of the personnel sought by government and industry to assess hazards and resources.



Report of the Research Briefing Panel on Cognitive Science and Artificial Intelligence

Research Briefing Panel on Cognitive Science and Artificial Intelligence

William K. Estes (*Co-chair*), Harvard University
Allen Newell (*Co-chair*), Carnegie-Mellon University
John R. Anderson, Carnegie-Mellon University
John Seely Brown, Xerox Palo Alto Research Center
Edward A. Feigenbaum, Stanford University
James Greeno, University of Pittsburgh
Patrick J. Hayes, University of Rochester
Earl Hunt, University of Washington, Seattle
Stephen M. Kosslyn, Harvard University
Mitchell Marcus, Bell Telephone Laboratories
Shimon Ullman, Massachusetts Institute of Technology

Staff

Paul S. Rosenbloom, *Rapporteur*, Carnegie-Mellon University
David A. Goslin, *Executive Director*, Commission on Behavioral and Social Sciences and Education
Sarah M. Streuli, *Administrative Secretary*, Commission on Behavioral and Social Sciences and Education
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

Report of the Research Briefing Panel on Cognitive Science and Artificial Intelligence

Start with a fundamental question that arose with the scientific world view: How can mind exist in a world totally governed by physical law? Shift that question, with the coming of experimental psychology: What laws and regularities govern human behavior? Convert that question, with the growing understanding of biology: How can the brain give rise to all the phenomena that we call mental? Expand that question, with the emergence of computer science: What is the general nature of intelligent action, independent of the device (brain, computer, etc.) by which it is implemented? All these questions reflect a single underlying great scientific mystery, on a par with understanding the evolution of the universe, the origin of life, or the nature of elementary particles.

Nothing short of a full history of Western intellectual endeavor can make sense of the crazy quilt of disciplines devoted to facets of this fundamental question, each in its own way. Cognitive science and artificial intelligence stand together in taking information processing as the central activity involved in intelligent behavior and in taking the framework of modern computer science as the foundation for understanding information-processing systems. Cognitive science reflects a concern with how humans process information, being an offshoot of cognitive psychology, linguistics, and philosophy. Artificial intelligence reflects a concern with how computers process information, being a part of computer science. The rapid growth of these disciplines in recent years is advancing our understanding of the nature of mind and the nature of intelligence on a scale that is proving revolutionary.

The research cited in this report reflects three main themes:

1. Answering the fundamental scientific question of the nature of intelligence and how it arises out of primitive cognitive functions. A subtheme, but one pregnant with hope, is the joining of this understanding with work in neuroscience, to make new progress in grounding the mechanisms of intelligence in the mechanisms of the brain.
2. Solving increasingly complex problems by constructing computer systems with the requisite intellectual, perceptual, and learning powers—making full use of the increasing speeds and memory being provided by hardware technology.
3. Developing human talent by combining

an understanding of how humans learn and perform cognitive tasks with effective intelligent systems to aid this development.

THE DOMAIN OF COGNITIVE SCIENCE AND ARTIFICIAL INTELLIGENCE

How can we characterize this broad field? Look first at what investigators are typically found to be doing. An investigator in artificial intelligence is constructing a computer program that can detect trends and constancies in data and induce physical laws. The program (aptly called *Bacon*) has rediscovered Kepler's third law of planetary motion and Galileo's law for the pendulum, among others. The investigator is trying to produce a program that can accomplish a task requiring kinds of reasoning hitherto presumed to be the province solely of human intelligence. He is also progressing toward discovery programs of greater generality through the evaluation and improvement of the approximate rules (heuristics) used to guide the program in its search for physical laws.

A cognitive psychologist is measuring the speed with which people can carry out mental operations on visual images—for example, rotating a figure in imagination—and observes that the time required to rotate a figure mentally is proportional to the angle of rotation, as specified externally. The investigator is testing hypotheses about the form in which information is represented in images and working toward a model capable of simulating the way mental operations on images enter into reasoning and problem solving.

The scope of the combined field of cognitive science and artificial intelligence is much wider than suggested by these two illustrations. In one direction are formal investigations of the abstract properties of computational systems that are basic to all intelligent behavior. In another are experimental studies of human observers detecting signals of millisecond duration in noisy backgrounds. In yet another direction are analyses of verbal communications to find the rules whereby people resolve ambiguities in language.

The common core of these diverse enterprises may be taken to be the task of understanding intelligence and intelligent behavior. What might we expect of a general theory of intelligence? Most immediately, it should help us to understand how the constituents of intelligent function have evolved in the human mind, the limits they set on possible achievement, the potentialities that are open to improvement. At the same time, the theory should help us understand how computers can carry out human-like activities and the ways in which they may exceed human capabilities. More broadly, an adequate theory of intelligence should guide attacks on questions or issues that go beyond our ordinary experience. Considering the current searches by some physicists and astronomers for evidence of intelligent life elsewhere in the universe, would we recognize intelligence in organisms very different from ourselves? Is it possible to set any theoretical limits on the possible differences in level or quality of intelligent function? Is human intelligence subject to biological constraints that do not apply to artificial systems?

ORIGINS

Artificial intelligence has been a recognized field of study for only a little over 25 years. Its theoretical substrate was the development of ideas regarding mathematical logic and computation during the 1930s and 1940s. The rapid development of digital computers from 1950 on provided a challenge to get these machines to accomplish tasks that had previously been the sole province of human intelligence. The computer provided, for the first time, a mechanism whereby theories of intelligent performance could be rigorously defined, built, and tested—providing the feedback necessary for rapid progress. The first efforts were the well-known attempts to program computers to play intellectual games, notably chess. But as it became clear that intelligence was based on the ability to process symbols and that computers were general symbol manipulators rather than just high-speed numerical calculators, the scope of efforts rapidly broadened to include research into induction, problem solving, theorem proving in logic, language understanding, and vision.

Cognitive science emerged into public view in the late 1970s, as an interdisciplinary field to cover all the sciences adopting an information-processing approach.* It did not come out of thin air. Seventy-five years of work by experimental psychologists brought human cognition into the laboratory and developed models sufficient to enable meaningful measures of the speed of mental operations and the amounts of information stored in memory following a learning experience. However, the ability to address how people deal with problems much more complex than those of simple laboratory tasks came with two developments of the 1960s.

The first was a suitable conceptual framework for a science of cognition. Developments in computer science and artificial intelligence gave rise to the view of the human as an information processor—a special kind of processor perhaps, but still one sharing basic principles and constraints with computer-based information-processing systems. The computer made clear how to partition the staggering task of comprehending the human mind and brain in all their aspects into two more tractable components. The critical idea was the distinction between computer hardware and software. One can learn how a computer works by studying its hardware—its electrical circuitry, recording medium, and timing mechanisms. With little knowledge of such matters, one can master the software and become a highly skilled programmer.

In the study of human intelligence, the scientific objective is to construct a theoretical picture of the structures and processes of human cognitive function at the level corresponding to software—the theoretical notions referring to the way information becomes encoded and organized in memory, the cognitive operations available for accessing and transforming the information, and the functional principles governing the sequencing and combination of operations. Cognitive scientists continue to hope for the ultimate explanation of mental phenomena at lower levels of scientific analysis, but they have found out how to proceed toward the development of useful theories without waiting on developments in neuroscience. These cognitive theories are thoroughly mechanistic, even though the mechanisms are not yet related to how they are realized in the brain.

The second main source of concepts and methods for the new cognitive science came from advances in linguistics. The importance of language in human thought has always been recognized in a general way. But the beginnings of a clear understanding of what it means to know a language, and therefore the first possibility of discovering how languages are learned and used, flowed from developments in theoretical linguistics in the 1960s, in particular Chomsky's formulation of generative grammars—sets of abstract rules that produce exactly the set of grammatical sentences. Increasingly sophisticated and psychologically relevant theories of grammar have provided a framework for interpreting studies of language-based cognitive activity and posed in clear and researchable form the major problem of how the understanding of language is achieved and how language is acquired by the young.

FROM THEORY TO RESEARCH

Research in artificial intelligence and cognitive science is strongly theory driven.

---

*As such, artificial intelligence is included in cognitive science; we have kept them separate here to emphasize the joining of human and engineering disciplines.

Increasingly, empirically oriented research is aimed at finding evidence relative to general theories rather than being a process of accumulating empirical facts and attempting to induce laws or principles. Thus a look at theoretical issues concerning system design and function will lead directly to the principal lines of research within which we can identify special opportunities with a high yield from increased efforts.

SYSTEM DESIGN

In artificial intelligence, system design refers to the relation between the way computers are constructed and the kinds of programs that can be implemented on them. The key concept is the *architecture*, which is the structure of the machine that permits it to be programmed—thus, the structure that creates the distinction just discussed between hardware and software. Any particular part of a system can be implemented in either hardware or software, with the choice to be made on the basis of trade-offs involving implementation cost, speed of execution, frequency of use, functionality, and flexibility. A critical question is what mechanism must be included in the hardware in order for the system to be able to manifest intelligent capabilities. What mechanisms, if not built into the architecture, would make basic intelligent capabilities (such as learning and problem solving) too difficult or slow to be feasible?

We have come to understand that one major aspect of the architecture is that it provides for *symbolic* behavior. That is, it permits computers to have internal representations that *refer* to other things. The computer's symbols (which are basically its addresses and operation codes) seem special, referring only to data structures in its own memories and to its own information processes. Yet out of that mode of reference comes the ability to refer to things in the outside world—to objects, activities, events, and abstract concepts. The internal symbols access data structures plus programs that interpret the data structures as representing the external thing. These data structures and programs were constructed by processes that had information about the outside thing. Because of this we ourselves can program computers to deal with many things—anything to which we set our mind. But systems themselves can also construct such symbols and their meanings when given information about the external object. Thus, for the first time we see how symbols and symbolic behavior can be realized in a physical system.

The hypothesis that links artificial intelligence and human cognition together, providing the frame within which scientific inquiry proceeds, holds that all intelligence, including human intelligence, arises from the ability to use symbols. By symbols we mean the physical symbol systems, noted above, as first understood in computer science. Without this stipulation the hypothesis would constitute only another version of the long-held view that symbols are important in human mental life. But with the notions of symbol and symbol system firmly connected to the physical world via the realization in computers, the hypothesis anchors the study of all the higher processes of mind to the same scientific world of mechanism as all other natural sciences.

Much is known about the abstract structure of such symbol systems (so-called universal computational systems), so the linking hypothesis is not that somehow the human is like a computer but that both humans and computers embody these systems and gain, thereby, their ability for intelligence and information processing.

Whereas the architecture of a computer is determined by the designer, the architecture of the human cognitive system is given by nature, and discovering its properties presents a major scientific challenge. For instance, it is evident that a human does not carry out actions the way current computers do, by a simple interpretive cycle of fetching the next instruction and then executing it. Rather, the human is much more recognition driven, with the current situation (at each instant), including the goals and momentarily active memory, directly determining the next action to be taken. The human cognitive system seems to be highly general-purpose, taking on a particular shape only in the course of adapting to the demands of a specific task. But structural constraints—those aspects of the cognitive architecture that determine which intelligent functions are easy or difficult to perform—are clearly of major importance. We know the human brain has many highly developed special structures for specific functions. Resolving the nature of these constraints is a prime task for research in cognitive science.

System Function

In artificial intelligence the computer is made to take on intelligent function by programming it within the framework provided by the architecture. The endless diversity of intelligent behavior stems from the ability of an intelligent system to program itself—to create symbolic structures to guide its own future behavior. Programmability per se does not make evident how the intelligence is attained, even given a suitable architecture. What sort of programming leads to intelligent action? This too is now clear in outline. The two basic ingredients are search and knowledge.

Search arises by defining a space of possibilities large enough to contain the sought-for solution and then searching for the solution by generating the states of the space. Thus arises the familiar exponentially expanding search of potential future chess positions by chess programs as well as the search of possible isomers of a given chemical formula by the *Dendral* program in searching for an isomer that satisfies the mass spectrogram data of an unknown sample, thus discovering its structural formula. Indeed, all the basic methods of intelligence appear to involve search.

Knowledge is necessary to guide the search through the space—to avoid having to find the proverbial needle in a haystack. A prominent component of artificial intelligence systems is the memory organizations that hold the knowledge and permit finding the right bit of knowledge at the right time. For instance, much of the progress in the development of expert systems—programs capable of exhibiting expert-level performance in intellectual domains—has come from discovering that large amounts of how-to knowledge can be effectively represented by collections of *if-then* rules (productions), where each rule of behavior carries the conditions in which it is to apply (the if-part) as well as what it is supposed to do (the then-part). In contrast to how-to knowledge, bodies of factual knowledge are organized into networks of associations, so that access to part of the knowledge base provides the connections to obtain other relevant parts. These networks go beyond the simple concept of association. For example, they must be organized such that widely applicable knowledge is not redundantly stored with each specific object to which it applies. Knowledge about an automobile must be composed of what is specific to the car (its dents) plus knowledge about all cars of that make. But the latter is similarly composed of knowledge specific to the make-plus knowledge about cars in general. This progression continues up through the most general nature of physical objects, imposing a hierarchical as well as an associational structure on the knowledge. The discovery of these forms of representation and memory organization and their efficacy for wide classes of tasks constitutes important scientific knowledge about intelligent systems.

Beyond matters of organization is the question of content. For a computer system to solve problems about cars, it must know about cars—and not just the specific facts about Fords and Chevies but all the concepts that make up the domain, such as steering, carburetors, that motors heat up, what a motor is, what heat is, and so on. Artificial intelligence systems differ from more standard computer applications precisely in that the knowledge involved in the system is not highly circumscribed. By the same coin, our current artificial intelligence systems, despite much progress, are still woefully limited in their knowledge compared to what general intelligence requires. Obtaining the knowledge to incorporate in systems is itself a major scientific task. This is especially true of expertise. One of the interesting results to come out of the work on expert systems is that experts cannot explain in general to another person the knowledge behind their ability to make expert choices. Only when put into concrete choice situations can they evoke reasons for their particular choice, and these reasons must then be integrated into a total organization (as in the *if-then* rules above).

Intelligence is shaped to an important extent by the properties of the languages and representations that it employs. Human intelligence is intimately involved with natural language (e.g., English), and in this respect it contrasts with the programming-language representations available in computers. The extent of the resulting constraints on human thinking and the degree to which they can be incorporated in computers are not yet fully understood. Human thought is also involved to an important, and perhaps not fully appreciated, extent with imagery. A new line of research in cognitive science is the investigation of models that may enable us to understand the "language" of the imagery system.

MAJOR RESEARCH AREAS

With the background above, we can now turn to specific areas of research.

ARCHITECTURES

It should be apparent from the discussion so far that architecture is a key concept for cognitive science and artificial intelligence, bringing in its train a large number of significant notions. Research into architectures is an active and important part of the field.

Computer Architectures

The most active research on architectures is in artificial intelligence rather than cognitive science. This effort is concerned with discovering the best architectures to support powerful problem-solving and learning mechanisms. Not only must computers be created with immense power (artificial intelligence supercomputers), but many organizational issues in the architecture are still to be discovered and understood. For example, how should memory be structured and accessed so that contact can be made between a novel situation and stored knowledge and problem-solving methods? How can new information be integrated into existing memory structures so that the system can learn through experience? Research into such topics translates directly into more powerful applied systems. It also provides a fund of knowledge about possible mechanisms upon which theories of human cognition can be built.

Massive Parallelism

The study of massively parallel systems has recently emerged as a major focus of architectural research. Originally, of course, computers were primarily uniprocessors (and artificial intelligence research concentrated on aspects, such as problem solving, that were primarily serial). But for some time computer science has been devoting major attention to how to exploit parallel computation, driven by the increasing difficulty in developing ever faster circuits and attracted by the emergence of integrated-circuit technologies that favor parallelism. There are many flavors of parallelism, varying with the number of concurrent processors, the intimacy of communicative coupling between the processors (measured by the instructions performed between significant interactions), and the volume of data flowing between them.

Massively parallel architectures are designed to have many simple processors (millions), with each one dedicated to a conceptual entity represented in the system. This contrasts with current architectures (including parallel ones) that dedicate a segment of memory to each conceptual entity but share processing capacity by shipping each piece of data to a processor for its moment of modification. The immediate research problem is learning how to compute effectively with these million-fold parallel architectures. However, the research vista that structures of this new class open up for cognitive science and artificial intelligence is more than just the hope of increased power. This class of architectures makes new contact with biological computation, which is also massively parallel. Indeed, the work is being deliberately guided by biological systems, especially the visual system.

Human Cognitive Architecture

Whereas in computer science and artificial intelligence, research on architectures takes the form of designing architectures and investigating their properties, in cognitive science the approach must be to hypothesize possible architectures and discover whether their implications for cognitive processes accord with empirical observation. The task is formidable because the university of plausible hypotheses is immense, and properties of architectures are only remotely connected to observable behavior. A major step toward making the enterprise realistic was taken in the 1970s. The key notion is deriving models from cognitive theories that, when realized as running computer programs, not only embody the theory but also perform the cognitive functions. A (soundly conceived) model of how a human learns can itself learn, and a model of how a human solves problems can itself solve problems. This operational property provides an important link in the chain of inferences from theory to observation, allowing behavioral consequences of complex architectural mechanisms (and combinations of them) to be directly determined. Comprehensive models of the human cognitive architecture have not yet been attempted, but several useful models for substantial subdomains, such as verbal memory, reading, imagery, and skill acquisition, have been put together. This trend toward integration is fundamental to making basic progress.

Sensory Information Processing

Most of our knowledge of the external world comes to us through either our eyes or ears — vision and speech. Hence, understanding the workings of these systems is a prime scientific problem. It is also critical for technological reasons. Vision is necessary for autonomous robots or similar devices that could simulate human performance in industrial settings or in the exploration of space. Speech is necessary to provide the ultimate natural means of communication between humans and machines.

Visual Perception

The end product of the analysis of a visual scene is subjectively familiar. We recognize the objects present in the environment and their spatial relations to each other and are able to compute distances in a way that permits us to navigate; also, the visual experience leaves a residue that can somehow be revived in the form of visual images, which enter importantly into reasoning about natural phenomena. The very earliest stage of the process leading to these consequences is well understood. Light reflected from the visual scene activates a mosaic of receptor cells on the surface of the retina at the back of the eye. But how information about objects in three-dimensional space is extracted from the two-dimensional mosaic of stimulated points and transmitted through the nervous system over nerves that carry information only by virtue of rates of firing of nerve impulses presents a formidable problem. It is somewhat paradoxical that the retinal image contains an enormous amount of information (for example, an aerial photo is typically digitized into about 70 million bits) yet underdetermines the three-dimensional scene that gave rise to it. Thus a device that could build a representation of a scene in the manner of human vision must incorporate knowledge other than the momentary visual input. Discovering precisely what knowledge will suffice and what computation must be performed on the input to yield a three-dimensional representation is a prime task for cognitive science.

The early stages of visual processing produce representations of basic information, such as contours, texture, motion, and depth, that provide the basis for determining shape. All these pose very hard scientific problems, both to discover what the mammalian visual systems do and to produce computational systems with equivalent performance. On none of these aspects are we close to attaining human levels of performance. This is not for want of practical incentive; high-quality systems for any of these would be extremely useful in artificial vision systems and would feed directly into the development of robotics. Indeed, substantial artificial intelligence research has already occurred on them all.

The current lead problem of this kind is that of stereo vision, namely determining just how the brain combines the images of a visual scene coming from the two eyes and uses small differences between them to construct a representation of the scene in depth. However, each image contains a very large number of elements, and it is a most difficult problem to construct search algorithms to find the element of one image that corresponds to any given element of the other. Some progress has been achieved by what are termed relaxation methods, but stereo algorithms take inordinately long to run on available computers. Recent developments in computer hardware technology, using very large scale integrated circuits, now provide the opportunity for significant advance. Using very large numbers of parallel computations, the three-dimensional field can be generated efficiently from the two-dimensional mosaic of retinal receptor elements. There are high hopes that in the next few years this problem will be solved, and we will obtain computational performance in the realm of human performance. This would represent a signal achievement, opening the way to an assault on the other problems of this type.

Because the exploration of low-level vision algorithms is being driven by what we know about psychophysics and neuroscience, this is the area in which cognitive science and artificial intelligence are in closest touch with neuroscience and where the prospects for a highly profitable bridge are best.

Speech Recognition

Since the 1950s, recognizing speech has been a prime target both of speech scientists and of technologists trying to develop recognition machines that would enable easy communication between human and computer. Speech recognition might seem a simpler problem than vision because its input is only one dimensional, not two, but speech has additional complications stemming from the complex encoding produced by the vocal apparatus. What appears to the native listener as a discrete sequence of words is, viewed acoustically, a continuous signal without breaks. Each small sound interval is the result of several contiguous segments of the speech it purports to represent—posing a genuine decoding task. Much of the progress to date has come only with the drastic simplification to recognizing separated words rather than continuous speech—the current commercial market is built upon this reduction. The recognition of words in normal continuous speech remains a major open problem.

In speech, as in vision, ambiguity of perception can be resolved by using knowledge and search heuristics from the cognitive realm. Efforts to bring these resources to bear on continuous-speech recognition programs in the 1970s yielded substantial progress—but only to the point of handling limited vocabularies (1000 words) in systems with strong syntactic and semantic constraints. Recent empirical work, however, has demonstrated that human listeners can recognize continuous speech efficiently using only low-level knowledge of words and phonology, together with knowledge of how basic speech sounds are realized in the acoustic wave form. Implementing these results in artificial intelligence programs offers promise of a new burst of progress.

NATURAL LANGUAGE

The understanding of natural language has come to be one of the central problems of cognitive science. (By *natural language*, we mean language as learned and used in ordinary life, as distinguished from the formal languages of mathematics and computing.) One reason is the important role of language in human thought and communication, another the need for machine comprehension and generation of language in order to permit easy and effective communication of humans with computers.

Learning Natural Language

In dramatic contrast with the difficulty of constructing a computer program to understand unrestricted natural language is the speed and apparent ease with which young children learn to cope with language. Preschool children outperform any artificial intelligence program, and the average six-year-old builds new vocabulary at a rate of more than 500 words per month, at the same time perfecting, at an automatic and unconscious level, mastery of a complex of grammatical rules that enable comprehension and production of an infinite variety of sentences. Moreover, language comprehension requires a semantic analysis of the message conveyed by the language, and learning new concepts often parallels the acquisition of linguistic descriptions for these concepts.

Before we can expect to cope with the full complexity of this learning process, we must accomplish the more basic tasks of investigating the induction mechanisms and knowledge organization principles that give rise to human learning. Significant recent results on learnability, together with some new developments in linguistics having to do with lexically based grammars and their incorporation in models of syntactic processing, are beginning to yield an idea of some of the properties of a reasonable theory of language acquisition. However, the investigation of learning methods is an endeavor that transcends natural language to encompass the acquisition of new skills, concepts, inference rules, and knowledge representations.

Understanding Natural Language

Investigating how adult human beings (or machines) comprehend and produce natural language need not wait on the development of a theory of language learning—and, in fact, has progressed faster. Such research bears on the possibility of easy communication with machines via typed or spoken commands and on the way information is gained from text or discourse in education learning.

The study of natural language understanding can be viewed from three mutually supportive objectives:

1. The identification of linguistic universals and the codification of grammatical systems into principled theories providing a formal characterization of language. This objective is associated primarily with the field of linguistics but clearly affects the automatization of lan-

guage comprehension by computers and the understanding of possible mechanisms that may underlie human communicative abilities.

2. The simulation of human linguistic performance with a view toward understanding human communication, the cognitive processes that support linguistic communication, and the inference system required to extract the intended meaning from the language. This approach is typically associated with cognitive science, and it affects the design and construction of practical language communication systems in major ways.
3. The development of computer systems that can interpret language in limited but increasingly sophisticated domains. This objective is the central artificial intelligence concern with natural language—making computers able to communicate with humans untrained in the intricacies of any particular formal language.

Natural language understanding systems produce internal representations of the information conveyed by the system. The most useful internal representations are *canonical* (recognizing that there are many different ways of expressing the same thought in any language), *semantically rich* (encoding information inferred as well as information explicitly found in the surface form), and *unambiguous* (unlike the natural language itself). These international representations are then applied to a variety of tasks such as text comprehension, data-base query, command interfaces, story interpretation, etc.

Artificial intelligence systems have gradually expanded the capabilities for language comprehension by machine:

- *Lunar* was one of the first truly integrated systems that performed grammatical analysis and subsequent semantic interpretation to answer questions posed by geologists about rock samples obtained from the Apollo-II mission. (For example, "What is the average concentration of aluminum in the high alkali samples?")
- *Lifer/Ladder* was the first semipractical system used in data-base query. The final version allowed some tolerance for user errors such as misspellings or the use of occasional sentence fragments.
- *Margie, Eli, Sam, Pam,* and *Politics* were based on a scheme of encoding semantic relations called conceptual dependency. Rather than striving to provide user interfaces, these systems attempted to emulate human comprehension and focused on performing inferences about the intents, plans, and possible outcomes of the speaker in a dialog or of characters in stories. This work demonstrated how language comprehension is an integral component of the total human cognition rather than a separable skill. The same mechanisms used for solving problems, planning actions, and recognizing situations come to play in interpreting and generating language.

Additionally, theoretical work in the structure of dialogs, such as speech-act theory, dialog focus, and belief spaces, has paralleled the implementation of working-language interpreters. Moreover, recent advances in linguistics, such as Lexical Functional Grammar, are unifying previous purely syntactic approaches (such as Chomsky's transformational grammar) with semantic interpretation methods. These parallel theoretical developments, together with recent work on robust parsing (i.e., understanding ungrammatical language) and advances in generating language from conceptual structures, promise to yield a new generation of more flexible and powerful language understanding systems. Such systems are starting to be prototyped as unified practical language interfaces to expert systems, data bases, operating systems, and other computational facilities.

Neither the artificial systems nor the models of human comprehension yet come close to a satisfactory treatment of full language understanding. Nonetheless, this area is one in which returns for investment in research are sharply on the upswing.

Intelligent Systems

The recent highly publicized advances in artificial intelligence in developing so-called expert systems reflect progress in understanding how substantial knowledge can be used to perform tasks. They have used realistic tasks, such as medical diagnosis, petroleum prospecting, and industrial assembly, and enough highly specific knowledge to almost entirely eliminate search. For example, *Mycin*, one of the classical expert systems, performs diagnosis for particular realms of bacterial infections (e.g., blood diseases) and does so at the level of a medical expert. *Mycin* is a collection of several hundred rules, each of which expresses a bit of diagnostic knowledge and each of which actively calls itself forth whenever it is relevant to advancing the diagnostic task. Each rule is applicable to only a few situations, but in the aggregate they provide a mosaic of coverage that produces, in effect, a field of responses that guides the program in all situations. This knowledge, extracted from expert physicians, represents codified know-how. Though strands of medical science weave through it, just as in all clinical medicine, it is really the accumulated wisdom-applied-to-action of the experts, not their ability to reason or understand.

Systems with Full Expertise

The progress into programming with lots of knowledge, rather than with lots of search (or with highly repetitive algorithms, as in scientific computation), although it has already opened up interesting domains of application, simply sets the stage for the next round of important research. One immediate issue, for instance, is how to weld together extensive knowledge and extensive search in a single system in which the spaces of possibilities to be searched can be created and evoked wherever appropriate.

Another major issue is how to provide such knowledge-intensive programs with a deeper understanding of the substantive domain of their expertise, so they can reason their way to conclusions and can explain why its advice is sound. (Present systems do provide some explanation, but it is relatively shallow, in accord with the superficial character of their knowledge; for instance, *Mycin* can present the rules used in a particular diagnosis to explain what it did.) A substantial body of active research on the nature of such basic understanding is concerned with examining how human experts understand their domains of expertise. The cognitive science research into the nature of expertise is in close contact with the more applied work on expert systems being pursued within artificial intelligence. As a result, we expect the current rather shallow expert systems to metamorphose into much more capable systems —and more scientifically satisfying examples of intelligent agents.

Learning

Studying simultaneously how intelligent systems work and how they acquire their capabilities is so large a task that over their early decades both artificial intelligence and cognitive science resorted to fractionation. In each case the first focus of effort was on the static aspects of information processing.

Artificial intelligence has so far made its major gains by concentrating on performance—on what is needed by way of representation, however, and search to accomplish tasks of varying degrees of difficulty and specification. Nonetheless, the problem of learning is a bottleneck for both practical and theoretical reasons. Expert systems with millions of knowledge rules can be generated only by learning, and in truly intelligent systems learning must go on concomitantly with all problem solving. So until learning is understood, we are walking on one leg. Sparked by these needs, and a better grasp of the theoretical and computational tools required to approach the problem, a resurgence of work on learning is under way.

In cognitive science we have just come through a period in which attention has been focused almost exclusively on the end products of learning—the way acquired information is represented and organized in memory. With some significant results in hand, we are seeing a better balance, with the emergence of studies in concept learning and skill learning. The former line is elucidating how people learn both formal and natural categories (for example, logical or mathematical as compared to biological categories or colors). In the latter line, new experimental approaches in conjunction with models are yielding insight into the way extremely high levels of skill are attained in some areas (for example, the performance of mnemonists and "lightning calculators").

The increasingly close intertwining of the artificial intelligence and cognitive science approaches raises prospects of significant advances in our understanding of learning in all of its various guises. The results can be expected to spill over to the full field of artificial intelligence, as learning mechanisms are incorporated into performance programs in myriad ways, and should amplify the applicability of cognitive science to problems of education.

SCIENCE AND MATHEMATICS EDUCATION

Educating people in technical subjects is a good example of how work from many parts of cognitive science and artificial intelligence comes together. It is both a currently active area of research that shows great promise for making fundamental advances and an area that is capable of making solid contributions to the nation's education problems. Current distress with early science and mathematics education is due in part to social and economic factors but also in part to aspects of learning and instruction where contributions might be sought from cognitive research and theory. The issues and potential contributions are somewhat different in two major problem areas.

Tutorial Systems

One of the roots of difficulties at higher levels of education is generally believed to lie in our continuing inability to produce acceptable skill at reading, spelling, elementary arithmetic, and the like in the full population of normal school children. A similar problem is becoming acute in the training or retraining of adults in the basic knowledge and skills needed for employment in jobs with increasingly high technical demands.

From the standpoint of cognitive theory, what might be a source of our persisting lack of success at producing basic skills in the full population of normal schoolchildren? A firm generalization arising from research on learning of skills is that acquiring significant competence in domains of any difficulty requires large amounts of guided practice—much more than is provided for most pupils in schools or most adults who try to prepare for new technical jobs.

The combination of advances in cognitive theory and basic advances in computer technology (providing interactive workstations of substantial power at increasingly modest costs) makes possible significant inroads into the problems posed by technical education. What can be expected of computer-implemented instruction at this juncture? An intelligent tutoring system that can provide genuine help in educating a student in some well-understood domain, such as mathematics or science, must provide several components.

1. A powerful model of the task domain, so it can itself solve problems in that domain.
2. A detailed model of the student's cur-

rent level of competence, encompassing both partial and erroneous competence as well as perfect competence.

3. Principles for interpreting the student's behavior, so as to be able to infer the student's knowledge and difficulties.
4. Principles for interacting with the student, so as to lead the student to a higher level of competence.

Experimental tutoring systems already exist for significant aspects of geometry, algebra, and arithmetic, and others are currently under development for some aspects of computer programming and elementary physics. For instance, a system dubbed *Euclid* works with a student who is learning to prove theorems in geometry—querying the student, evaluating progress, and providing advice.

The work in tutoring also feeds back and contributes to basic understanding of cognitive science and artificial intelligence. Recall *Mycin*, the artificial intelligence expert system mentioned earlier. The *Mycin* analysis of its task domain of diagnosis made it a candidate for an automated tutor capable of instructing medical students. An attempt to construct such an intelligent tutor showed that *Mycin*'s knowledge was too purely diagnostic know-how and that *Mycin*'s problem-solving style was too different from appropriate human strategies. A revision of the *Mycin* system (called *Neomycin*) attempts to provide a much more cognitively relevant treatment of the task, thus leading back to relevant research on the expert system side.

As artificial intelligence provides expert systems in an increasingly wide variety of tasks, and as the models of students' performance and learning proliferate, the potential exists in this area for a substantial contribution to education in science, mathematics, and technology. Additional elements not yet technically feasible, such as natural language interaction, should soon become possible. Progress on natural language interfaces for tutoring systems is expected to move at a more rapid pace than general research on language comprehension because the constrained nature of the domains to be taught allows simpler, though less general, techniques to be used.

Mental Models and the Problem of Transfer

To have an adequate supply of college-age students ready to move toward careers in science or engineering, effective education in mathematics and science must have begun much earlier. But it is a hard fact that these subjects are regarded as difficult, and very large numbers of children, perhaps including some of the most talented, are turned off by their early experiences. Furthermore, results of cognitive research suggest that faulty starts may be very costly. Inappropriate modes of thinking instilled at an early age, or during early experience with new material at any age, can be very difficult to modify later.

When solving problems in their field, experts mentally construct *qualitative* models of their tasks and reason within these models by applying knowledge and search. For example, when trying to solve elementary physics problems, expert physicists reason qualitatively about the situation until it becomes clear what calculations to make. They construct a diagram of the physical situation and reason to the variables that determine the motion. In contrast, naive problem solvers attempt to solve such problems by going directly to reasoning mechanically with the equations.

The ability to do qualitative reasoning goes hand in hand with expertise, but qualitative models per se do not necessarily lead to successful problem solving. The physicist's qualitative models are useful because they are built around the concepts of force and energy. When naive problem solvers try to bring their own expertise from daily life to bear on physics problems, their qualitative models look more like the earlier pre-Newtonian impetus models (a ball shot from a circular pipe continues to curve, although gradually straightening out). These naive theories provide a framework within which novices interpret demonstrations intended to teach them new facts and begin to learn to solve quantitative problems. Unless these informal theories are understood by the teacher and painstakingly supplanted by new habits of qualitative reasoning about natural phenomena, they persist as a barrier to learning. Thus, even after conventional instruction in new physical, or other scientific, concepts, students who do well on textbook examples often prove incapable of applying the laws or formulas they have mastered at a formal level to interpret actual physical events. The proper sequencing of experiences with qualitative and quantitative reasoning is, consequently, a critical aspect of early science and mathematics education.

The difficulty of producing transfer of learning from educational experiences to new problems has been a bugaboo since the earliest days of educational psychology. The first major research results discredited the prevailing idea of "formal discipline," the notion that general memory and problem-solving abilities could be strengthened by such measures as instruction in geometry and the classics. But nothing was left in its place except demonstrations of the extreme specificity of most learning to its original context. (In World War II, prospective aerial gunners were given intensive practice on skeet ranges—the sole discernible result being skill at breaking clay pigeons.)

It has become clear that for materials or tasks of any complexity, certainly including all forms of science and mathematics learning, a key to transfer is conceptual understanding. Important recent advances in cognitive science include empirical methods for assessing a person's qualitative understanding of principles in a domain (e.g., motion of objects on surfaces in kinematics) and characterization of the mental models, akin in many respects to scientific models, that individuals form when learning about causal relations in such domains. One intensive new line of research attacks the problem of just how the mental models and modes of problem representation of the truly expert solver of problems in specific areas of science or engineering differ from those of the less competent. Beyond the immediate objective of attaining theories of competence comparable to those now available in linguistics is the longer-term goal of tracing the developmental path from the mental model of the novice to that of the expert. A by-product of this effort is a principled basis for adaptive competency testing—on-line interpretation of a student's problem-solving efforts that yields an assessment of understanding. The methodology is advancing rapidly and seems capable of extrapolation from simple mathematics to such complex practical tasks as nuclear plant operation.

NEEDS AND OPPORTUNITIES

Realizing the emerging opportunities for important theoretical and practical advances depends on the degree to which some important needs for support are met. Currently, much more support is available to computer science and artificial intelligence than to cognitive science. This imbalance stems directly from their perceived roles as being hard and soft sciences, respectively. Yet it should be clear from this entire report that the two disciplines closely interlock and contribute to each other at every level and on every scientific problem. Cognitive science especially needs resources to permit it to contribute as a full partner.

The situation in basic artificial intelligence research has a different character. Surrounded by the sudden expansion of applied artificial intelligence, with its rash of new venture-capitalized firms and industrial laboratories, it is hard for basic artificial intelligence research to compete for the talented scientists. Even much basic research is seen by those doing it as very close to real application; hence, bright young scientists are tempted to move toward the commercial sector. From a broad social perspective this somewhat frenetic activity is merely the concomitant of obtaining the social payoff from science. But it tends to deplete the next generation of basic science. The only solution we know is the standard one of attempting to make the basic research situation sufficiently attractive to counter the dispersive factors.

Two specific factors have tended to dampen progress and may do so more seriously in the future. One is a shortage of computing power, and the other is a dwindling supply of new talent. These affect both cognitive science and artificial intelligence, although from somewhat different sides, as noted above.

Computers are the laboratories of artificial intelligence research and major components of laboratories in cognitive science. Unfortunately, both new theoretical ideas and new possibilities of application are outrunning the computing power needed to implement them. Many well-trained cognitive psychologists and other behavioral scientists are unable to move into newly active lines of research in cognitive science because they do not have access to the computing power and languages required for the development of theoretically significant systems. Even in the research groups where the most intensive efforts in artificial intelligence are going on, what appear to be significant opportunities for new theoretical advances are having to wait until some way is found to give investigators access to substantially increased computing power.

Two distinct kinds of computing power are needed. One is the networked powerful personal computer capable of substantial list processing. Adequate research environments in artificial intelligence now require essentially one machine per active investigator (which includes a substantial fraction of graduate students). The other kind is special hardware systems or artificial intelligence supercomputers (not to be confused with numerical supercomputers, such as the Cray and Cyber machines). These systems are all experimental at the present time and can be built only in computer engineering environments and often only in cooperation with industry.

A factor that is not yet critical but that might seriously constrict the advance of research within a few years is an insufficient supply of new talent. Theoretically significant research in cognitive science and artificial intelligence requires extensive technical training in mathematics, psychology, and computer science. Young people are unlikely to go through the rigors of this training unless there are prospects for research careers. The dwindling opportunities for faculty appointments in cognitive science at research universities is by no means compensated by the increasing opportunities in industry, for the balance of support for applied versus basic research is quite different in the two cases. If we want many talented young people to embark on the road to expertise in cognitive science, it would need to become known that a commitment to the support of basic research in this area on the part of the government and universities is sufficient to provide reasonable prospects of research careers.

The needs of this field for carefully engineered support are well balanced by the emerging opportunities for significant theoretical and practical advances. A number of promising possibilities are apparent in the research areas surveyed in this report.

1. The development of comprehensive and executable architectures for cognition.
2. The understanding of massively parallel computation, both at the level of computer architecture and at the level of models for cognitive processes.
3. The development of effective programs for computer recognition of visual scenes, language, and speech, drawing on the resources of both artificial intelligence and cognitive science.
4. The research that attempts to bridge the gap between both artificial intelligence and cognitive science and neuroscience.

5. The application of what is being learned about high-level human reasoning and processes of skill and knowledge acquisition to the development of a new level of expert systems in artificial intelligence.
6. The development of computer-based intelligence tutoring systems, which can play an important role in educating and reeducating our people for a high-technology society.

However, we do not wish to place too much weight on specific lines of research. Cognitive science and artificial intelligence differ from various predecessors most importantly in the emphasis on general problems of fundamental theoretical significance. Though a full theory of intelligence is still remote, the insight that both intelligent machines and their human counterparts can be seen as instances of the class of symbol-manipulating systems has yielded some general results and opened the way for more.

Broadly viewed, research in both cognitive science and artificial intelligence is showing that the acquisition and organization of information are basic aspects of intelligence in both humans and machines. A salient and pervasive characteristic of human beings, distinguishing them sharply from lower organisms and special-purpose computers, is a tendency to acquire information that goes far beyond current task demands and to organize the accumulating knowledge so that it can be accessed and used in unforeseeable future problem situations. Building this feature effectively into computer systems can be seen as the next major advance in the power of intelligent systems. Cultivating it during human cognitive development may be a key to materially increasing human problem-solving abilities.

An important offshoot of research on artificial intelligence is its clarification of the problems and its development of methods pertaining to the systematization and codification of the enormous accumulation of knowledge in technical domains. It is coming to be widely believed that information processing will be the heart of the next industrial revolution. (Witness the scale of the Japanese Fifth Generation Computer project, centered on very large scale machines for artificial intelligence.) As occurred earlier for energy and the physical sciences, and later for medicine and the biological sciences, only massive basic research efforts can prepare us to keep up with this fast-moving worldwide development.

But machine intelligence alone will not suffice. Human intelligence must comprehend the emerging problems and the machines that aid in their solution. Thus, new challenges face our already heavily burdened education system. Cognitive science may be able to provide the broadened theoretical framework within which significant progress can occur and artificial intelligence major components of the needed new educational technology.

Report of the Research Briefing Panel on Immunology

Research Briefing Panel on Immunology

Hugh O'Neill McDevitt (*Chairman*), Stanford University School of Medicine
Barry R. Bloom, Albert Einstein College of Medicine
John J. Farrar, Hoffmann-La Roche, Inc.
C. Garrison Fathman, Stanford University School of Medicine
Mark I. Greene, Harvard Medical School
Leroy E. Hood, California Institute of Technology
Kimishige Ishizaka, The Johns Hopkins University School of Medicine
Richard Lerner, Scripps Clinic and Research Foundation, La Jolla, California
Philippa Marrack, National Jewish Hospital and Research Center, Denver, Colorado
Stuart F. Schlossman, Dana-Farber Cancer Institute, Boston, Massachusetts
Jonathan W. Uhr, The University of Texas Health Science Center at Dallas
Irving Weissman, Stanford University School of Medicine

Staff

C. Garrison Fathan, *Rapporteur*, Stanford University School of Medicine
Barbara Filner, *Associate Director*, Division of Health Sciences Policy, Institute of Medicine
Linda Peugh, *Administrative Secretary*
Talitha A. Shipp, *Secretary*
Gail Wilson, *Secretary*
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

Report of the Research Briefing Panel on Immunology

INTRODUCTION

Four questions have long intrigued immunologists:

1. How does the immune system have the ability (or the potential ability) to construct an antibody to virtually any foreign material—thousands and thousands of distinct foreign molecules, some not even found in nature but constructed in the laboratory?
2. How does the immune system recognize "self" and accept it while rejecting that which is foreign?
3. How does the immune system's memory work (the ability to respond more quickly and more strongly when challenged with a foreign material it has "seen" previously)?
4. How do the several cell types involved in the immune system communicate so that an integrated, regulated response effectively removes foreign material?

In the past two decades, immunologists have gained substantial understanding about the biochemical structure of the antibody molecule and the processes that enable the genetic system to generate such a great diversity of antibodies. They have learned of two major cell types—B cells and T cells—and have inquired deeply into the regulation of functions of the B cell. The T cell has been more elusive. However, recent advances in basic knowledge and some powerful new technologies have created major new scientific opportunities to identify the role of the T cells involved in regulation of immune responses and to understand the surface molecules (receptors) and secreted molecules (lymphocyte hormones) that mediate necessary interactions.

This new knowledge is leading to such a sufficiently clear understanding of the workings of the immune system that it can now be manipulated toward quite precise experimental and clinical goals. We can study the molecular mechanisms not only of the immune response but also of other basic biological processes, particularly those entailing cell interactions mediated by surface molecules or by minute amounts of protein signals.

Such recently acquired knowledge and techniques also have tremendous potential for the improvement of human health, for making possible effective new means of diagnosis, therapy, and disease prevention. In fact, the potential already is moving toward realization at the level of clinical experimentation.

The technological tools that have allowed these opportunities to develop include monoclonal antibodies, cloning of genes, cloning of immunocompetent cells, and automated and highly sensitive sequence analysis and synthesis of nucleic acids and peptides.

The major research opportunities for the near term are the T cell receptor and lymphocyte hormones (lymphokines), both activators and suppressors.

This paper will explain immune responses, the power of the technological tools, and the reasons why the T cell receptor and lymphocyte hormones would be significant and fruitful areas for scientific inquiry. The paper concludes with a discussion of the additional opportunities that would be offered by this new knowledge—opportunities to advance basic knowledge of the immune system, its responses and regulation of those responses; opportunities to advance knowledge of other basic biological processes and of disease processes; and opportunities to exploit this knowledge with new approaches to the prevention, diagnosis, and treatment of an array of diseases, such as autoimmune disorders, cancer, genetic disorders of blood, infectious diseases, and virtually any health problem for which organ transplantation offers a solution.

IMMUNE RESPONSES

When the body is challenged by a foreign substance (antigen) that is extracellular, such as bacteria, certain cells of the immune system respond with proteins (antibodies) that recognize and bind to the antigen. That response, called "humoral immunity," is mediated by B cells, which secrete a large number of antibodies into the blood. The antibodies mobilize many other immune system activities to rid the body of the foreign substance. A different kind of immune response is at work in intracellular infections, allergic reactions, or rejection of transplanted organs. This entails "cell-mediated immunity," a function of T cells.

Both B and T cells are lymphocytes originating in bone marrow. Each B or T cell recognizes a specific antigen and can synthesize a unique receptor that recognizes the antigen. Upon appropriate contact with the antigen, the B or T cells are stimulated to divide. They produce a "clone" of identical cells; cells in the clone produce identical receptor molecules and thus have identical specificity for antigen.

One of the great research successes of the past two decades has been the answer to the first question we posed: How does the immune system construct unique antibodies against a seemingly infinite array of foreign substances? The determination of the amino acid sequences of the protein chains in antibodies* and parallel genetic studies revealed that part of the immune cell's DNA is rearranged in order to provide the requisite antibody diversity. In other words, genetic endowment is not rigid and the precise structure for each antibody molecule is not inherited. Gene parts are provided, but they are sorted and resorted to form many new combinations.

This surprising discovery opened the door to a truly revolutionary view of all genes—one that encompasses control of expression during growth and development by rearrangement of a portion of DNA from one site to another along the length of a chromosome.

---

*A typical antibody molecule consists of two pairs of protein chains—two identical heavy (long) chains and two identical light (short) chains. Each chain, whether heavy or light, has what are called "constant" regions and "variable" regions. The constant regions are the same for a large number of antibody molecules, even though their specificity varies. This part of an antibody molecule confers properties held in common, such as helping to trigger inflammatory responses. The variable region, which provides specificity for antigens, is what makes each antibody clone unique.

Recent research has shown that this is indeed true for a number of genes, not only those determining antibody structure.

MAJOR TECHNOLOGICAL GAINS

Recombinant DNA technology and relatively stable and homogeneous B cell lines were essential for the advances in knowledge of B cells. With new technology, the time has now arrived to investigate T cells, their receptors, and their regulation. The relevant technological achievements are the production of monoclonal antibodies, the cloning of genes, and the cloning of immunocompetent cells. Advances in sequence analysis and chemical synthesis of nucleic acids and proteins offer additional promise for the future.

MONOCLONAL ANTIBODIES

In 1975, Kohler and Milstein showed that certain malignant mouse cells that produce antibody molecules (myeloma cells) could be fused with normal antibody-producing B cells. The resultant "hybridoma" cells can be grown continuously and indefinitely as cell cultures that synthesize large amounts of antibody molecules. The antibody produced by such a clone of hybridoma cells is called "monoclonal antibody" and has the important property of having a single specificity—all of the antibody molecules are identical.

Because of unlimited possibilities in producing hybridomas, there is the potential to produce antibodies against a seemingly endless variety of important antigens. There is the technical hurdle of finding a clone with the desired specificity for antigen, but considerable ingenuity has been exercised in screening and selecting for clones of interest.

The discovery of hybridomas and monoclonal antibodies launched a new era in immunologic research and in the application of immunologic assays to basic and clinical problems. One research application has been in the identification of T cell types and elucidation of their biological functions. Other research applications cross a broad spectrum of biological areas.

CLONING OF GENES

The genetic information specifying the structure of protein molecules, such as enzymes and antibodies, is encoded in a cell's DNA. It is translated into a working copy—messenger RNA—for use by the protein synthesizing apparatus of the cell. Advances in genetic engineering (recombinant DNA technology) have enabled scientists to isolate messenger RNA from cloned cell populations with defined characteristics, such as the synthesis of a particular protein, make a DNA copy of the genetic blueprint inherent in the RNA, and then incorporate that DNA copy into the DNA of a virus. Monoclonal antibodies can then be used to detect those viruses that have "expressed" the protein of interest during their infection and reproduction in appropriate cells. Once the virus and then its DNA are isolated, the gene of interest can be purified.

Thus, by using recombinant DNA technology and monoclonal antibodies, one can isolate in very few steps the genes that encode proteins of biological and clinical interest. Antibody-directed gene cloning should lead to the rapid isolation of almost any gene, provided the researcher can define a functional polypeptide (protein-like molecule) and the cell that produces it. These immunologic and recombinant DNA technologies should continue to produce a rich diversity of reagents for identification and isolation of biological macromolecules of scientific, clinical, and commercial interest.

In particular, gene cloning should enable immunologists to study in detail the function of genes expressed selectively in particular cells of the immune system. Included would be antigen receptors located on the surface of T cells, other receptor systems that govern interactions between lymphocytes and other cells, and the hormones that regulate the growth and differentiation of the immune system's cells.

Cloning of Immunocompetent Cells

Recent discoveries have made possible three different approaches to the long-term growth of T cell clones in cell culture. One method adapts the hybridoma technique first developed for B cells and fuses malignant T cells with normal T cells. A second approach depends on the addition of "growth factors" to the cell culture medium. A third approach involves repeated stimulation of the cultured T cells with a particular antigen and serial reculturing.

These three approaches have set the stage for the identification and isolation of a variety of cell types and secreted cell products involved in regulatory events mediated by immunocompetent T cells. Preliminary experiments using similar approaches suggest that it will be possible to readily clone other sets of immunocompetent cells, including B cells and antibody-presenting cells. This suggests the eventual possibility of constructing an integrated in vitro network of the cells that govern immune responses.

Chemically Synthesized Antigens

Proteins, which consist of long chains of amino acids folded into a three-dimensional structure, trigger immune responses and thus are active as antigens. Certain peptides, which are short chains of amino acids, also can act as antigens. Some chemically synthesized peptides have been found to mimic the relevant structure in natural antigens and to elicit immune responses.

What once took the efforts of several laboratories over a period of years (e.g., determination of the amino acid sequence of a single protein chain) can now be done overnight or in a period of days with automated machinery. The technologies that enable rapid and accurate analysis of the amino acid sequence of proteins and the chemical synthesis of peptides offer opportunities to probe the basis for antigenicity. In more applied terms, these technologies also offer opportunities to produce exceedingly safe new vaccines that completely lack an infectious agent. Chemical synthesis also provides the possibility of tailoring an antigen to be a more potent immunologic stimulant. New vaccines against hepatitis, polio, rabies, and herpes viruses are being designed using this approach.

For the future, there is no reason to restrict research to peptide mimics of antigens. Organic chemicals should be considered as well. If we learn the rules by which proteins can be mimicked by organic compounds, we can think in terms of entirely new classes of drugs. This would have applications not only in vaccine development but also in the treatment of hormone and enzyme disorders.

RECEPTORS AND SIGNALS THAT REGULATE THE IMMUNE SYSTEM

The immune system has two main attributes: recognition of self and nonself and immunologic memory. Both recognition and memory are a function of the receptors on the surfaces of immunologically competent lymphocytes and of the protein products elaborated by such cells.

The receptors on B cells that enable them to recognize and respond to foreign materials have been known for some time. They are immunoglobulin (antibody) molecules, closely related to the antibody molecules these cells secrete after stimulation. The protein structure of these molecules and the genes encoding them have been among the most rewarding areas of immunologic research in recent years.

The T Cell Receptor

Characterization of the antigen-specific receptors on T cells has been elusive. There are two major reasons for the difficulty: (1) T cells, unlike B cells, do not secrete soluble receptor materials, and (2) the receptors on T cells bind only to antigen when it is associated with proteins on the cell surface membrane, not to free antigen. This second phenomenon, identified by pioneering work in the mid-1970s, has been termed "restricted recognition." The cell membrane proteins involved in T cell recognition are the products of a collection of genes that constitute a locus called the major histocompatibility complex (MHC). Restricted recognition of antigen by the T cell suggests that its responses will be directed only at cell-surface-associated antigens. A cytotoxic T cell, for example, will kill a virus-infected cell that bears virus antigens and MHC products on its surface but will not interact with circulating free virus. (Free virus will be removed by antibody that is the cell-free product of B cells.)

For the immunologist the property of restricted recognition makes it difficult to describe the nature of the antigen-MHC protein complex recognized by T cell receptors. Indeed, it is still not known whether the T cell has two different types of receptors, one for antigen and one for MHC, or a single receptor that recognizes some sort of aggregate of the two molecules. The phenomenon of restricted recognition means that binding to antigen or to MHC cannot be used as an enrichment technique for the isolation of T cell receptors for a particular antigen. The situation has prevented our understanding how one-half of the immune system (T cells) recognizes antigen and regulates the function of the other half (B cells).

Two of the new techniques outlined above, however, have contributed to the recent isolation of T cell receptors for antigen-plus-MHC by several laboratories. One technique depends on B cell hybridomas; the other employs cloned T cell lines or T cell hybridomas with receptors of a single specificity for antigen-plus-MHC. Several laboratories have shown that mice immunized with cloned T cell lines or hybridomas produce antibodies that either block or mimic recognition of antigen-plus-MHC by these T cell lines. In every case published to date, the antibodies are specific for the immunizing clone and do not interfere with antigen-plus-MHC recognition by other T cell clones, even those with similar specificities. Apparently it is the variable portion of the T cell receptor that is being recognized by the relevant antibodies.

The discovery of the T cell receptor for antigen-plus-MHC is so recent—and so much remains to be done—that it is difficult to predict exactly what advances will be realized. Certainly it will lead to the isolation and understanding of the genes encoding these products and to knowledge of the structure of the antigen-plus-MHC product that the T cell receptor recognizes. Additionally, it is known that certain cell surface molecules are associated with T cell receptors for antigen-plus-MHC and contribute to the overall interaction between T cells and their targets. These molecules—LFA-1, OKT4, and OKT8—may have genetic variants that affect their function. Isolation of these molecules and receptors will enable us to understand their function and perhaps to control their action.

Basic understanding of some disease processes also can be foreseen as possible extensions of this research. Many inherited diseases in humans are genetically linked to the MHC, including some so-called autoimmune diseases, such as juvenile-onset diabetes, rheumatoid arthritis, and multiple sclerosis. Because recognition of antigens by T cells is associated with recognition of MHC products, it is thought that diseases of this type are caused by some sort of abnormal T cell response to antigen-plus-MHC, which leads to cross-reaction with a self-antigen-plus-MHC. Understanding how T cell receptors recognize antigen-plus-MHC should lead to clarification of autoimmune phenomena and perhaps to development of methods of circumventing them.

Additional potential therapeutic applications can be envisioned. If the structure of the T cell receptor is known, and monoclonal antibodies can be raised against unique determinants of it, it should be possible to delete selected clones of T cells. For example, such techniques could eliminate clones of leukemic T cells or cells involved in organ graft rejection. Such an approach has the advantage of leaving the rest of a patient's immune system intact.

Lymphocyte Hormones

The recent discovery of a new class of molecules elaborated by lymphocytes has led to new developments in understanding cell regulation in the control of immunity. Under the proper stimulus, lymphocytes can produce molecules that stimulate other cell types to divide and replicate. These molecules are called "interleukins," which are a type of lymphocyte hormone. One interleukin active in tumor cells has been defined at the gene level, has been molecularly cloned, and has been shown to promote lymphocyte division. This linkage of immunology to cellular and molecular biology should provide valuable insight into the regulation of cell growth and cell division.

With interleukins, immunologists have been able to undertake cellular cloning of T cells. The clones of T cells grown under the influence of these molecules can now give insight into the molecular organization of the recognition system of these uniquely specific cells. Furthermore, it has been found that clones of lymphocytes obtained from mice exposed to tumors and grown under the influence of interleukin-2 can retard and eradicate a similar tumor in other mice. It is anticipated that such interleukins (of which three distinct types have already been defined) will have application to tumor therapies. Interleukin-3 has been shown to promote the division of certain blood cells, suggesting that their growth may be regulated in part by this type of molecule. It is anticipated that interleukin-3 will facilitate the study of blood production by the body and lead to therapeutic modalities for some blood diseases.

Another area of study relates to the expected discovery of other interleukins whose growth-producing activities are restricted to other cells of discrete lineages. Although discovery of interleukins capable of stimulating activities of lymphocytes and other blood cells is clearly a major advance, it also may be possible to identify new genes or to molecularly alter already defined interleukin genes so as to make their hormone products capable of stimulating preferentially one type of cell whose absence needs to be corrected. Furthermore, these interleukins may help provide insight into signals and activities important in the growth of normal and malignant cells. They also may lead to a new family of pharmaceuticals with broad biological activities.

In contrast to the activities that promote cell growth and activity, another class of lymphocyte hormones that suppress cellular activities has been discovered. One type of lymphocyte, termed suppressor T cells, produces molecules that inhibit the development and expression of immunity. These molecules, which still require extensive biochemical definition, are thought to be polypeptides that operate in a sequence of cellular interactions to inhibit both B cell and T cell immune responses.

Areas of study for suppressor hormones relate primarily to the ways in which suppressor cells become activated by antigens, to the biochemistry of the molecule produced, and to the action of such cellular products in the inhibition of cellular processes. Previous research has been hampered by the very low concentrations of some of the suppressor hormones—insufficient amounts of hormone were available for biochemical and biological analysis. But the detailed study of the structure of suppressor lymphocyte hormones will undoubtedly be facilitated by B and T cell hybridoma technology. If both hybridomas that secrete homogeneous products and monoclonal antibodies for these products are developed, it should be possible to clarify the structure of the hormones. It also will be possible to determine the genes that encode these inhibitory substances and to study their organization.

The further development of human suppressor hybridomas is important. Since allergic diseases are caused by antibodies in a class called IgE, selective suppression of IgE antibody formation would be the fundamental treatment for these diseases. Human T cell hybridomas that produce IgE suppressor factors or that could be used for cloning of genes controlling such factors would provide a new strategy to control allergic diseases. Human hybridomas capable of elaborating nonspecific suppressor molecules have recently been defined. Thus, the development of more specific human suppressor hybridomas may have therapeutic application to transplantation and inflammation in addition to allergic responses.

Some mouse hybridomas secrete specific substances that inhibit inflammatory responses similar to those that occur in many pathological states of human beings. The development of human suppressor substances may prove useful in the treatment of autoimmune, allergic, and connective tissue disorders, in which abnormal humoral and cellular responses contribute to morbidity.

Suppressor cells may inhibit myeloma growth, suggesting that lymphoid dysplasias may benefit from new types of suppressive therapy. However, in other solid tumor models, suppressor cells can become activated and inhibit eradication of tumors. Thus, it is likely that another area of development will include strategies to bypass or regulate suppressor activity, which could lead to therapies for human cancers.

APPLICATIONS OF NEW IMMUNOLOGIC KNOWLEDGE

These advances in knowledge of the immune system will play a crucial role in furthering our understanding of basic biological processes and will have medical applications as well. The basic biological processes studied certainly will include the mechanisms and control of immune responses but also will include virtually any biological process involving cell-surface alterations, cell-surface interactions, or cellular communication via peptides present in very low concentrations. We have indicated already the additional important applications to basic research involving isolation of genes.

Basic Research

Over the next decade there will be a rapid accumulation of antibodies reactive with many cell types and cell products. Already there are monoclonal antibodies that distinguish the two major classes of lymphocytes in the immune system, T and B cells, and that recognize subsets of these cells specialized to perform different functions. Antibodies will be developed to react with the new lymphocyte hormones and their receptors on the target cells. The ability to distinguish the multiple cell types and cell products participating in immune responses will help to elucidate the systems that control these responses.

It is likely that monoclonal antibodies will be generated that react with cells of many organ systems and even with subsets of cells in those organ systems. These monoclonal antibodies to various cell-surface antigens would constitute a library of molecules that distinguish cells in a precise way that was hitherto impossible. It is expected that the surface antigens will help identify different phases of function, different stages of maturation, or different lineages of differentiation. This identification will help elucidate the processes involved in control of diverse biological processes.

Medical Applications

The potential diagnostic, therapeutic, and preventive applications of this new knowledge in immunology are too numerous to discuss here in detail. New approaches are apparent for autoimmune diseases, cancer, and genetic defects in the blood-forming system, for example. Management of infectious diseases, especially via vaccines, and virtually any disorder necessitating organ transplantation also will benefit.

Examples of the readily foreseeable medical uses of monoclonal antibodies, lymphocyte hormones, and recombinant DNA technology are provided by juvenile-onset (insulin-dependent) diabetes mellitus and cancer.

Juvenile-Onset Diabetes

This disease affects 0.5 percent of all children in the United States. It leads to lifelong dependency on injected insulin and to a number of severe long-term complications. Recent research indicates that juvenile-onset diabetes is due to an autoimmune response—antibodies against their own insulin-producing cells are made and the cells are destroyed. It is possible to detect an autoimmune response months or even years before actual development of clinical diabetes.

We already know that most children with this disorder have two particular transplantation antigen types, DR3 and DR4. However, screening is not yet feasible because fully 35 percent of the normal population has these two antigen types. Further biochemical and molecular genetic analysis should identify subtypes that would enable prospective identification of a much smaller population at risk. Screening would then be feasible.

Isolation of the target molecules on the insulin-producing cells would open the way to specifically suppressing this autoimmune response. Use of this target molecule or of human monoclonal antibodies directed against the transplantation antigens involved in the development of autoimmunity might permit therapy that suppresses the autoimmune responses.

Thus, it should be possible within a few years to identify in advance a relatively small percentage of the population at risk for this disorder. These children would be tested on a regular basis for development of the autoimmune response, and treatment to prevent the development of diabetes could be initiated after early signs of autoimmunity.

In the case of juvenile-onset diabetes, monoclonal antibodies seem to offer the best opportunity for identifying, treating, and understanding this autoimmune disorder. For other autoimmune diseases in which there is too much or too little of a lymphocyte hormone, hormone analogs or purified hormones offer possibilities of restoring normal immune regulation.

Cancer

Monoclonal antibodies that are completely pure, have a single specificity, and show high affinity should aid in the discovery of tumor-associated antigens. Antibodies to these antigens offer opportunities for diagnosing tumors that are not precisely identifiable by conventional techniques and could even become the basis for new cancer taxonomies. These antibodies could be used to stain frozen sections, to stain cells dissociated from biopsy material, for tumor imaging, and to analyze blood using flow cytometry. Monoclonal antibodies to tumor-associated antigens also could be used to develop very sensitive radioimmunoassays that will detect these antigens in various body fluids.

One of the major problems will be to determine the distribution of any tumor-associated antigens on normal tissues. If the antigens are not unique to a given tumor, it may nonetheless be possible to use a panel of monoclonal antibodies tailored to assay for particular cancers. The results should allow earlier and more accurate diagnoses.

Monoclonal antibodies also offer new approaches to cancer treatment. They could be used for direct antitumor activity or for targeting delivery of toxic chemicals to tumor cells. For example, there are many highly toxic proteins produced by bacteria or plants—the plant product ricin, for example. A single molecule is sufficient to inactivate protein synthesis and to cause death of the cell in which it is present. Experimental studies indicate that toxins coupled to a tumor-reactive antibody are highly useful in ridding bone marrow of tumor cells. (In the autologous bone marrow rescue approach to the treatment of cancer, the patient's own "cleansed" marrow is transplanted back after massive doses of antineoplastic drugs or irradiation to kill tumors.)

The importance of bone marrow transplantation must be emphasized. In addition to therapy for certain cancers, it is the most obvious means for gene replacement therapy. In genetic deficiency diseases, especially those involving blood cell abnormalities, specific genes could be transferred into bone marrow stem cells, which would then reconstitute the normal functions of the recipient. But even when the donor and host are matched at the major histocompatibility complex, graft versus host disease has a rather high mortality, and this has impeded significantly the use of bone marrow transplantation. Fortunately, animal experiments and preliminary clinical trials indicate that antibody-conjugated toxins directed aginst T cells are capable of "cleansing" bone marrow of those cells that are responsible for the graft versus host disease.

Infectious Diseases

We close this discussion of medical applications with a brief mention of infectious diseases. Recent developments make it possible to purify antigens to an unprecedented extent for use as vaccines. Several methods are feasible, including using monoclonal antibodies to help isolate and purify antigens from natural sources. (One procedure often produces a 5000-fold enrichment in a single step.) Monoclonal antibodies and recombinant DNA technology also allow the genetic engineering of vaccines, and chemical synthesis is possible as well. Yet another approach entails antibodies made against the active site of a specific antibody. Occasionally the anti-antibody mimics the antigenic structure that elicited the specific antibody and engenders an immune response.

An example in which these techniques have been spectacularly effective is the production of a vaccine against the infectious form of the malaria parasite. The sporozoite can be obtained only from the salivary gland of mosquitoes, which do not lend themselves to vaccine production. However, a protective antigen was identified by a monoclonal antibody, the genes of the parasite were cloned in *E. coli*, and recombinant clones expressing the malaria-specific protective antigens were prepared. The vaccine has been found to protect experimental animals.

GROWTH OF IMMUNOLOGY RESEARCH

The truly impressive accomplishments in this field were made possible by the previous national investment of research resources. Vigorous growth in the past five years can be seen in a number of measures. For example, between 1977 and 1981 investigator-initiated research grant applications (ROI) increased more rapidly in immunology than in other fields of biomedical research without compromising quality. (There was a 27 percent increase in the number of competing proposals reviewed by the four immunology study sections of the National Institutes of Health but only an 18 percent increase in other fields.) Furthermore, between 1977 and 1982 the number of new PhDs in immunology increased steadily each year, rising from 101 to 150. We believe that with adequate research resources the scientific opportunities presented in this briefing paper will sustain continued vigorous growth.

Report of the
Research Briefing Panel on
Computers in Design and Manufacturing

Research Briefing Panel on Computers in Design and Manufacturing

Michael Wozny (*Chairman*), Center for Interactive Computer Graphics, Rensselaer Polytechnic Institute
Thomas Binford, Computer Science Department, Stanford University
Robert Brown, Electro-Optical and Data Systems Group, Hughes Aircraft Company
David Grossman, Automation Research, IBM Corporation
Christoph W. Klomp, Computing Systems, Boeing Commercial Airplane Company
James F. Lardner, Deere and Company
Michel A. Melkanoff, School of Engineering and Applied Science, University of California, Los Angeles
Stuart Miller, Automation Systems Laboratory, General Electric Company
Frank Stoneking, Engineering and Manufacturing Computer Coordination, General Motors Corporation
Arthur R. Thomson, School of Engineering, Cleveland State University
Herbert Voelcker, College of Engineering and Applied Science, The University of Rochester

Staff

George H. Kuper, *Executive Director*, Manufacturing Studies Board
Peter Smeallie, *Staff Officer*
Delphine D. Glaze, *Administrative Secretary*
Frances M. Shaw, *Secretary*
Courtland Lewis, *Rapporteur*, Herndon, Virginia
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

Report of the Research Briefing Panel on Computers in Design and Manufacturing

SUMMARY

America's future growth in industrial productivity depends vitally on the health of the nation's research in computers for design and manufacturing. In spite of the impressive apparent growth, the panel notes that the central issue in this field is a pervasive lack of scientific knowledge. Universities have made major contributions in the past 10 years with very limited resources.

The first priority of this field is to build a genuine classical research community through increased government support of basic research. The panel identified the following five key research areas that will significantly impact the field:

- geometric modeling and analysis
- human-computer interface for design and manufacturing
- knowledge-based or expert technology
- information management
- manufacturing and computer devices

The second priority is to build our university educational programs in an effort to strengthen America's technical excellence for the long term.

INTRODUCTION

Although the use of computers in design and manufacturing is relatively new, the rapid growth of computer technology and the promise of improved productivity and quality have drastically changed the way we can design and manufacture products. Clearly, issues associated with the application of computers in the factory workplace transcend technology and impact industrial organization, management, and business strategy. Computers will have a profound impact on the work force at all levels, including managers, professionals, and production workers. This report, however, is limited to technical research opportunities for increasing the use of computers that must be explored if the United States is to maintain a competitive edge in design and manufacturing.

RESEARCH GOAL

At the risk of oversimplifying, we can think of design and manufacturing together with marketing as three inseparable parts of a production cycle from conception to delivery. Marketing provides information to engineering and design on the product required.

Engineering designs the product and transfers product-definition data to manufacturing. The planning function within manufacturing transforms the product-definition data, e.g., the geometry, into process-definition data—the instructions for building the product—and then transfers this information to the factory floor. Computer-based planning and control systems directly schedule and track the manufacturing process as well as assure controlling the product's quality with regard to its design specifications.

Research on computers in design and manufacturing should lead to an understanding of the principles necessary to automate completely the creation, analysis, transmission, and management of all product-definition, process-definition, and related business data. This will ultimately take the form of a computer-integrated system. A corollary of this goal is to automate the physical processes corresponding to those data.

Background

University research has contributed significantly to the advancement of this field. For example, the modern concept of computer-aided manufacturing (CAM) had its start in 1949-50 at the Massachusetts Institute of Technology under an Air Force contract. Because of the need to machine aircraft components with complex geometries accurately and reliably, the project led directly to the development of numerically controlled (NC) machine tools. An NC machine uses numeric or symbolic instructions to perform a manufacturing function such as milling or grinding. A subsequent separate effort developed the Automatically Programmed Tool language, called APT, a computer language still used for programming NC machine tools.

During the early 1970s, university research efforts in the U.S. and several other countries created the basis for solid geometric modeling, a complete representation of an object that permits any well-defined geometric property to be calculated automatically. This concept removed a major roadblock to flexible automation and has tremendous near-term potential to revolutionize the industry. The U.S. university research was funded primarily by the National Science Foundation.

Focus on Mechanical Technologies

There exist two distinct technology areas where computers have been used effectively in design and manufacturing. One technology encompasses microelectronic chip design and fabrication. The other technology is associated with mechanical parts such as machines. Both technologies are very important. Although many of the generic manufacturing concepts, such as planning and scheduling, presented in this report apply to both technologies, the specific research needs for microelectronic design and process modeling are very different from the primarily geometric needs for mechanical design.

This report concentrates on the mechanical technologies because they represent an extremely broad range of diverse manufacturing applications whose mathematical models are not well understood. Mechanical technologies offer significant research opportunities, the results of which will benefit countless small industries. They also represent the fastest growing segment in the field of computer-aided design and computer-aided manufacturing (CAD/CAM).

ASSESSMENT OF THE FIELD

Computers allow us to routinely handle design situations whose complexity is far beyond the reach of an unaided human mind. Furthermore, the level of complexity that can be handled increases with each successive generation of computers.

For example, the computer contributed significantly to the management and design of a modern airplane such as the Boeing 767. The upper or lower wing panel on a Boeing 767 aircraft has stringers attached to it with

18,000 rivets each, the location of which is extremely critical. The wing panel is defined in the CAD/CAM system in a parametric form. More than 1,000 pieces of input information are needed to specify the exact panel location on the wing. Wing panel assembly had been a problem because it was impossible to accurately visualize the results until the stringer was essentially attached, a problem compounded by changes due to engineering error. This entire assembly was completed on the Boeing 767 using CAD/CAM. The computerized assembly allowed visualization of the entire panel prior to engineering release. It also greatly reduced error and improved fit-up so well that on the first panel assembly the actual assembly labor hours required were 72 percent less than originally estimated. In another example, the use of CAD/CAM reduced the time needed to fit the wing joint attachment from two weeks on the Boeing 747 to 18 hours on the Boeing 757.

Computers in Design

The computer extends the designer's capability in several ways. First, because the computer organizes and handles time-consuming and repetitive operations, it allows the designer to concentrate on more complex design situations. Second, it allows the designer to analyze complex problems faster and more completely. Finally, the computer allows the designer to share more information sooner with people in the company who need that design information (purchasers, tool and die designers, manufacturing engineers, process planners, etc.). These capabilities result in improved product quality, decreased design-cycle time, and designs virtually impossible to achieve manually.

The following examples illustrate these points. Using CAD, designers at General Motors were able to reduce the total weight of an average automobile by 25 percent and at the same time improve the overall structural integrity of the car. CAD reduced the time required to solve complex stress equations, allowing designers to increase the number of iterations and evaluate more completely the structural changes when metal was removed.

Using computers, GM engineers designing combustion chambers in automotive engines were able to interface a solid geometric model of the combustion chamber to a thermodynamic combustion analysis program, fire the spark, and evaluate how the flame propagates through the chamber geometry. For the first time, a designer was able to evaluate the effect of alternative geometric configurations on combustion without extensive mathematical support to set up the proper boundary conditions. Solid modeling allowed the designers to study analytical problems in a more direct and efficient way than was possible previously.

CAD systems are being used extensively in industry. Although most applications today are in two-dimensional drafting, sophisticated, but limited, three-dimensional surface and solid modeling systems are now on the market. The geometric modeling portions of these systems interface to analysis capabilities by allowing interactive problems to be undertaken. These problems are set up for finite element programs like stress analysis, for kinematics programs like mechanical linkage design, and for NC tool path generation. There is an economic trade-off between two-dimensional and three-dimensional systems, however. Three-dimensional systems require about four times more computer power than the two-dimensional systems. Companies are reluctant to acquire a three-dimensional system because they do not have the trained personnel to use it effectively. Although most CAD vendors are supplying color graphics terminals, their application appears to be for color line drawings rather than the generation of realistic color-shaded images.

Current trends in CAD are toward smaller, more powerful minicomputers. These systems have their own powerful built-in computers for stand-alone operations and can be connected to large mainframe computers for data base access or large analytical needs. Called intelligent workstations, such systems will begin to appear on engineers' desks in the near future.

Research opportunities in CAD are aimed at giving the designer more information about how his or her design will be manufactured. There are many examples in industry where different approaches to the design of the same part lead to different manufacturing procedures with widely varying costs. Future CAD systems will evaluate the merits of a design in progress and offer advice to the designer on the "best way" to proceed. Designers will also be able to access computer models of manufacturing processes for the purpose of evaluating the trade-offs of such contending processes on the manufacturability of the product.

An engineering drawing communicates the designer's ideas and requirements for a part to a machinist. The purpose of a dimension on a typical engineering drawing is to fix some distance or angle used to define an edge or surface of the part. The purpose of the corresponding "tolerance" is to give some acceptable deviation limits on that dimension. Tolerances are necessary because it is impossible to build parts with zero error or precise dimensions. Tolerances influence the machine and tool to be used, the required skill level of the operator, the expected part yield, and many other factors relating to the final cost and performance of the part.

An important step in the design of an assembly is conducting a tolerance "stack-up," or analysis, where potential clearance and interference problems that result from the buildup of tolerances are checked. Assemblies designed using CAD techniques have the advantage of an automated approach to conducting this tolerance analysis.

COMPUTERS IN MANUFACTURING

The major topics in this field are group technology, manufacturing planning and control systems, automated materials handling, computer-aided manufacturing, and robotics.

Group Technology

Group technology is the classification and coding of parts (material, size, shape, volume) and processes (lot size, machine tools, setup time) into families with similar characteristics. It enables the designer to evaluate designs to determine which parts already exist, which parts can be modified from existing parts, and which parts are truly new. Group technology facilitates the establishment of manufacturing cells (groups of machines coordinated to create a given part) for a family of parts, and provides a basis for computer-aided process planning. Direct benefits include a smaller number of parts in inventory (and in the data base), reduced time to design a new product, and increased use of capacity.

Group technology plays an important role in the planning function. Consequently, research is needed to develop models for manufacturing cells or hierarchies of cells that can be used to generate overall optimal scheduling strategies.

Manufacturing Planning and Control Systems

Manufacturing planning and control generally include the following functions: material resource planning (MRP), distribution resource planning, overall simulation and optimization, and shop floor data collection. These functions represent the management and business information needs of the entire enterprise.

The system developed over the last eight years by Ingersoll Milling Machine Company, a large discrete-parts job shop, illustrates the state of the art.* That system integrates a management and business information system with a CAD/CAM system. The system's master scheduler links the bill-of-materials and geometry files to coordinate dispatch, inventory, MRP, planning, and scheduling. High-level management decisions are automatically entered in the system, and each supply source (machine shop, purchasing, stockroom) receives only the appropriate information.

---

*The company received the 1982 LEAD award (Leadership and Excellence in the Application and Development of Computer-Integrated Manufacturing) from the Computer Automated Systems Association of the Society of Manufacturing Engineers.

This system has expanded dramatically the existing system's design and implementation capability and provided a major new role in strategic planning. Within four years the productivity of the system's design activity doubled. Today, 82 percent of the system's design time is spent in creative development and only 18 percent on routine maintenance (compared to 42 percent spent in creative development four years old).

The general trend in the area of manufacturing planning and control is toward the integration of more business planning and reporting functions. Although simulation techniques are used extensively to test scheduling scenarios, they cannot yet generate them. An important research problem is the refinement of large-scale optimization techniques to identify planning and scheduling strategies for large, hierarchical, flexible manufacturing systems. Other important topics include inventory management techniques to accomplish "just-in-time" objectives and the employment of knowledge-based expert technologies to ensure quality throughout the design, test, and manufacturing cycle.

Automated Materials Handling

The major automated materials handling systems consist of automated storage and retrieval systems, miniloaders, automated guided vehicle systems, and conventional materials parts storage and transport systems. Benefits include higher inventory record accuracy, reduced storage space requirements, increased labor productivity, and the automatic transmittal of material movement information from the production control system to materials handling equipment.

Computer-Aided Manufacturing

Computer-aided manufacturing generally refers to numerically controlled (NC) machining and other processes (including flexible manufacturing systems), computer-aided inspection and testing, automated assembly, and process control.

In the machining area, NC machines have been used for many years with proven benefits in the form of reduced time for machining and storage. The current trend in the industry is a logical extension of NC to 2 to 15 machine tools linked with automated materials handling and production scheduling—a machining cell. Such systems are commonly used to produce one or more families of parts. Benefits provided by these systems include increased production flexibility, reduced operating costs, and higher machinery utilization.

Computer-aided inspection is an area of intense research. The benefits include automatic collection and analysis of quality control data, creation of data bases to isolate production process problems, and correction of machinery problems before many faulty parts are produced.

A system developed at General Electric Corporate Research and Development for inspecting jet engine turbine blades represents the state of the art. The basic task for that system is to inspect turbine blade surface and subsurface geometry to exacting tolerances at high rates. The inspection system consists of three major subsystems: a subsystem to present and sense the part precisely, a signal-processing and image-processing subsystem to inspect the part, and an analysis and decision-making subsystem to accept or reject the part automatically.

In the electronics area, IBM has developed a robotics-based automated system to test the thousands of connections on the electronic boards of mainframe computers. This system is based on design data that describe the location of the connecting paths.

Automatic assembly systems today are generally built to perform one task. The trend in this area is toward flexible, electronically controlled systems, including assembly robots. It is possible to generate robot programs automatically for well-defined and regular geometric layouts such as circuit boards and cable harnesses.

Model-based CAM has the capability to generate computer programs automatically from geometry and other information, in order to control manufacturing processes and to simulate such processes for verification and evaluation. Automatic procedures would extract pertinent information from underlying models of various processes such as extruding, cutting, forming, casting, welding, painting, hardening, testing, assembly, and material flow. Mathematical bounds on the reliability and optimality of these models are also needed.

In primary processes, such as molding or forming, it is possible to derive mold geometry from part geometry. Modeling mold performance is still largely a research issue.

Logistics models involve material flow and process planning. From a network flow model it is possible to predict assembly-line behavior automatically. However, there is limited ability to derive optimal schedules. It is possible to derive process plans for the sequence of manufacturing steps required for axially symmetric parts.

Research is needed in modeling almost all processes. Models are also needed for sensor-mediated assembly and visual inspection, i.e., the use of sensors to control the process. Such models require better methods of modeling three-dimensional geometry and constraint relationships, including kinematics and tolerances. Methods are needed to permit fabrication planning from three-dimensional geometry.

Robotics

Robotics is an area of intense activity. The major industrial applications include spot welding, materials handling, finishing and coating, arc welding, and assembly. The two most promising application areas for future growth are materials handling, including loading and unloading operations, and assembly.

Current robots can offer very precise and repeatable position control. However, the availability of force sensing, the ability to control the force exerted by the robot's grasping mechanism, is very limited. Virtually no continuous force control is available.

Programming systems are somewhat incompatible with most standard operating systems and programming languages. They also lack features such as coexistence of interpreted and compiled code, configurability, and facilities for debugging, multitasking, multiprocessing, and networking. Smart (adaptive) programming systems are virtually nonexistent.

Current robot vision capabilities are two-dimensional binary and are generally restricted to well-separated parts. Development is under way on applications and systems for structured-light, three-dimensional systems (i.e., systems using light projected in coded form), which will greatly expand the capability.

Given current research, one can expect to see in the next several years general purpose, real-time distributed operating systems and programming languages. Robots will exhibit limited force control, assembly capabilities, and coordinated use of multiple arms. There will also be commercial, structured-light, three-dimensional systems, and limited model-based vision systems.

Methods are needed for linking robots and vision systems to CAM models to allow real-time use of such models. Research is needed on establishing a set of fundamental low-level commands to characterize sensor-mediated actions, such as force-controlled motion.

INTEGRATION OF THE TOTAL SYSTEM

The total integration of all computer-based systems that support design, manufacturing, and business functions is essential to realize the potential of the computer for increasing productivity. Although significant inroads have been made on a limited scale in specific industries, system integration on a large scale is an extremely complex and poorly understood problem.

The requirements of an integrated system include (1) analytical, operating, and control systems that communicate directly with each other; (2) systems able to generate (by creating, deriving, transforming, and interpreting data) the appropriate information necessary for the support and operation of the design and manufacturing functions; (3) the use of a common data base concept that can be physically distributed over many computers; (4) a responsive communications network; and (5) a master system to manage the whole.

Research is needed to understand the nature of the integration problem. We expect that simulation techniques will be useful in modeling integration concepts to help evaluate the anticipated effects, the reliability, the usefulness (cost/benefit), and the potential likelihood of achieving large-scale integration. Research on this problem is risky because of the comprehensive nature of the changes involved and the enormous investment required. Yet it carries the greatest potential for payoff.

Global Information Structures

We can think of existing research in design and manufacturing as creating the building blocks of a cathedral. We are putting the blocks in place to build the cathedral without actually knowing if the overall structure is going to stand up.

Currently, data carried in traditional industrial data bases is functionally divided to support engineering and manufacturing separately. The engineering data base supports geometry and engineering business data, such as scheduling and purchasing. The manufacturing data base supports the processing aspects of geometry, such as numerical control, and manufacturing business data such as planning and inventory control. In most cases these data are located in several physical data bases.

The data interfaces must be structured to support all the above industrial processes. The structure must be flexible enough to accommodate the information needs for different applications such as machining, scheduling, inventory control, estimating costs, analysis, or design. It is also desirable that the information be derived from the data without human intervention. Yet it must be general enough to permit the integration of current systems and still be extendable to future systems. Such an information structure will in effect become the master source for the definition of both product and process.

The nature of global information structures is not well understood today. Data flow concepts need to be developed and tested against a broad spectrum of applications.

Distributed Data Management and Networking

Applications requiring large data bases and involving different computers in a network create a severe communication problem because most data base management systems run on only one computer. Consequently, data communication problems arise because the data base management system in one computer cannot be used to query data in another computer.

Long-distance data communication between separate geographic sites is another typical problem in heterogeneous computer networks. Current services for the transmission of critical geometric and other data are slow, noisy, and costly. Efforts to establish standards are proceeding but are still unresolved.

Integration of the design and manufacturing of a product can occur only if the heterogeneous software and hardware systems support a common distributed data base. This distributed data base must have adequate recovery and backup procedures when simultaneous updates on different computers occur or when one or more computers in the network fail. It must have minimal redundancy of data while maintaining the most efficient access route and data traffic. Further, it must maintain audit trails of a large variety of simultaneous updates caused by design or manufacturing changes. Research is needed to study the data characteristics necessary to integrate the design and manufacturing systems within a distributed data base environment composed of heterogenous computers.

RESEARCH OPPORTUNITIES

The trends in the application of computers in design and manufacturing identified in the previous section suggest a number of research opportunities. Underlying these areas are a set of generic research topics that can be grouped into the five following areas: (1) modeling, (2) human-computer interface, (3) expert systems, (4) information management, and (5) devices. Research in these areas will provide the critical building blocks needed to speed the application of computers in the factory.

MODELING

Geometric models embody the description of a product design. Consequently, they represent a basic link between CAD and CAM. Most existing CAD systems describe product geometry by means of potentially ambiguous "wireframe" representations, a collection of points and lines. With wireframe representations there is no way of determining automatically if a point was "inside" or "outside" the object. It is not possible to calculate mass properties (i.e., volume, inertia) in any automatic way. Therefore, automated analysis (and model-driven CAM) will require a complete and unambiguous representation of the geometry in three dimensions.

Geometric Modeling

The ultimate goal of geometry research is to develop a complete and unambiguous three-dimensional geometric representation for part and process modeling. This representation should include free-form (sculpted) surfaces, allow creation of geometry in a natural manner (be user friendly), and permit automatic analysis and process planning.

Current Status The theoretical foundations for computer-based solid geometry representations were laid through university research conducted primarily in the U.S. and U.K. during the 1970s. U.S. industry is just beginning to use solid modeling in production. The solid modelers that are commercially available all have shortcomings.

Industrial research that is focused on methods for improving existing techniques should lead in a few years to solid modelers acceptable to industry with increased capabilities in surface sculpting and blending. Performance should also increase, primarily because of new hardware and improved software.

Key Research Needs An important research problem is the integration of free-form surface generation capabilities into solid modelers. Several approaches currently being investigated appear promising, including the rational B/SPLINE representation and the generation of blending surfaces between a given set of primitive solids. The deformation of solids is another important problem area where very few inroads have been made. The ability to handle deformations, for example, has very important practical implications in modeling the extrusion process. One can identify many such process operations that need to be investigated.

Another research area involves tolerancing information for solids. Current mathematical and computational theories of mechanical tolerancing have significant gaps. Finally, there are many geometry problems such as surface-surface intersection, stable position of free-form solids, numerical accuracy of geometric models, and new geometric primitives that need to be investigated.

Analytic Modeling

There are many analyses needed in the course of a design that involve the creation of an approximate (discretized) geometric model and then the application of loads and material properties. Thermal and stress analyses using finite element techniques are two examples.

Two steps are required to perform these analyses automatically. First, we need an automatic discretization procedure that will provide some type of approximate form of the initial geometric object. Second, we need a way of ensuring beforehand that the numerical results of the analysis will be within a prescribed level of accuracy (adaptive analysis).

Current Status Research is needed in analysis methods that can be applied directly to a geometry configuration with associated boundary conditions in some automatic way. For example, research on automatically generated meshes for finite element analysis is a very active area. Industry-usable results should be forthcoming in the next several years.

Key Research Needs Adaptive analysis is the primary area of needed research. We need the capability of applying an analysis procedure such as stress analysis directly to a geometric model with given boundary conditions and automatically obtaining an answer to a specified level. This problem has a high research priority because it represents an important step in the automation process. Expert systems are needed to aid the designer in choosing an appropriate approximation model for each condition in the problem (e.g., shell model or beam model for different parts of an airplane).

Human-Computer Interface

The human-computer interface in CAD is the link between a design engineer and the CAD system. A user-friendly interface will guide the designer through complex design tasks in a direct and effective manner. This is very important, especially in the early design decisions that have a major impact on the total design-production process. The development of such an interface requires an understanding of the creative design process. Ideally, the system will act as a natural extension of the designer. Friendly interfaces are especially important in CAM, where users may be less familiar with computing techniques.

Current Status Graphics systems are now very sophisticated, with the ability to create color-shaded, three-dimensional images from a design data base. However, most existing human interface capabilities are still relatively crude with very limited help capabilities to get out of trouble. In short, today's occasional user faces an imposing learning experience each time he or she sits down at the workstation. Graphics and engineering workstation vendors are beginning to pay attention to the relevant human factors and human engineering issues, but the effort is still very limited.

Key Research Needs Research is needed to understand better how an engineer conceptualizes, since it is in the early stages of design that current CAD/CAM is least effective. Modeling cognitive processes associated with design need to be explored. Research is needed in developing "smart" CAD systems that use expert (i.e., knowledge-based) concepts to guide the engineer interactively through a design situation by continually evaluating his or her performance and issuing guidance in terms of the individual's perceived experience base.

EXPERT SYSTEMS

Expert systems consist of a body of knowledge and a mechanism for interpreting this knowledge. The body of knowledge is divided into facts about the problem and heuristics or rules that control the use of knowledge to solve problems in a particular domain.

In the last few years, expert systems have become the most visible and fastest growing branch of artificial intelligence. Their objective is to capture the knowledge of an expert in a particular area, represent it in a modular, expandable structure, and transfer it to other users. To accomplish this goal, it is necessary to address issues of knowledge acquisition, knowledge representation, inference mechanisms, control strategies, user interface, and finally how to deal with uncertainty

Current Status Expert systems development, confined during the past decade to academic laboratories, is now becoming commercially viable, partly because of the development of well-understood methods for knowledge-based programming and partly because of advances in microelectronics. Current examples of expert systems include MOLGEN, which interactively aids molecular geneticists in planning DNA-manipulation experiments; VM (Ventilator Management), which gives real-time advice for the management of patients undergoing mechanical ventilation in an intensive care unit; PROSPECTOR, which advises when and where to drill for ore; and DELTA, a production rule-based system for diesel electric locomotive repair. Expert systems have not yet been applied in any real sense to solve manufacturing problems.

Key Research Needs Applied research is needed to gain some experience in building knowledge bases for selected manufacturing processes, especially in planning functions. Research is needed to sort out the issues of "design for manufacturability" before undertaking the task of building one. The concept of geometric reasoning is very complex; we need to understand how to characterize manufacturing "reasoning" such as "find a symmetrical object in the data base." Finally, we need to examine closely the issues or experiences involved in developing knowledge bases for making approximations in complex analysis problems.

INFORMATION MANAGEMENT

This area concerns the management of design and manufacturing data so that users and systems can use the data easily and efficiently.

A data base is an organized collection of discrete data on a given subject. A data base is made accessible and useful by means of a data base management system (DBMS). This management framework permits users to both query and update the data base in order to derive comprehensive lists of data having certain characteristics in common.

Current Status Data base management systems are commercially available for a wide range of applications. Current research is producing considerable advances in DBMS capabilities, including improved performance, modeling capabilities, and user interfaces. Research on distributed data bases is a critical area. However, there are features of the manufacturing environment that render current data bases and DBMSs inadequate. We do not know how to make data base systems that efficiently handle manufacturing requirements in a distributed environment.

Key Research Needs Research is needed to organize geometric data in a way that allows "geometric reasoning," i.e., to optimize the data extraction operations in response to continually developing and unanticipated user queries. For example, consider requesting a list of parts where the parts possess a certain symmetry, regardless of the sequence of geometric steps used to construct the parts originally. We may not know beforehand that symmetry is important and therefore would not want to assign a symmetry attribute. We want the system to "reason" its way to the correct information.

Other research areas include the following:

- A basic study into the nature of the integration problem. This could lay the groundwork for integrated experimentation on a larger scale.
- The development of data characterization requirements for the integration of systems in a distributed data base environment.
- The investigation of the information structure required to support the so-called neutral data base or master product definition.
- The development of algorithms to support proper recovery and backup procedures of distributed data bases in heterogeneous computer environments.
- The use of simulation techniques to help model integration to understand better the anticipated effects, reliability, usefulness (cost benefits), and the overall potential to achieve actual integration.

DEVICES

This section deals with issues at the device or workstation level as opposed to systems level. The basic technology of robot sensors and actuators must be advanced significantly in order to achieve the needed order-of-magnitude improvement in precision, speed, and dexterity. Component modularity involving new micromechanical, electronic substrate-based sensors and controllers is required for the flexibility needed to create unique manufacturing systems. The ability to integrate and coordinate comprehensive strategies for sensing and control is also very important.

These objectives imply the need for special-purpose computer architectures and VLSI (very large scale integrated) chip technology. Such chips will also be essential for handling the growing computational workload in automated CAD and CAM modeling and analysis operations. Chips for handling specialized geometric constructions, inference interpretations, and analysis processor calculations will dictate requirements for special-purpose microelectronic chips in the near future.

Current Status Currently, we have a heterogeneous collection of sensors and actuators and no uniform way of integrating these components into a coordinated manufacturing system. For example, vision systems are important not only for robots but also for all kinds of inspection and control functions.

Array processors for special-purpose analysis calculations and display processor chips for graphics are becoming more widely used. They are only the beginning of a trend toward fast hardware.

Key Research Needs Research is needed to develop smart sensors, including visual and tactile sensing. The investigation of modularity issues in sensors and actuators is also important. Finally, the development of requirements for special-purpose processors to implement advanced algorithms in analysis, geometry, expert systems, network communications, and control strategies are needed. The research issues here cut across other fields in a significant way. Collaboration will be required. This research is important because the ultimate precision and speed of our manufacturing will be dependent on the inherent capabilities of these devices.

THE ROLE OF UNIVERSITIES, INDUSTRY, AND GOVERNMENT

Universities, industry, and government traditionally have been the key institutions in manufacturing technology research. Each has its own method of addressing research problems, and each has unique limitations and strengths.

Support Mechanisms for Research

In universities, research tends to be small scale in nature, where specific, well-formulated problems are studied in a formal manner. Universities have been reluctant to grapple with the larger problems of integration, partly because of inadequate interdisciplinary knowledge (few faculty trained to address such issues) and because of an uncertainty about the scientific issues. Hardware and facilities are costly—usually beyond the means of a university budget. Even when hardware is donated by industry, the operational costs are substantial. In addition, many graduate students in computer-related fields are attracted by industry, sometimes before completion of their degree program. This effectively limits the formation of a cadre of qualified researchers. Government and other funding for university research in this field is insufficient to develop the critical mass of researchers needed to sustain significant progress.

Industrial companies have generally solved design and manufacturing problems piecemeal, addressing short-term objectives. Broad gaps in basic knowledge are increasingly apparent as computers are applied to factory problems that pose larger and more difficult integration hurdles. Compounding these organizational limitations are some situational problems.

1. It is difficult for industry to articulate its systems' integration problems. The issues vary greatly by company size and markets.
2. Manufacturing equipment is generally tied up for production purposes and unavailable for scientific experimentation.
3. The industrial application of research findings requires capital and personnel investments over a period of years, with accompanying high costs. Also required are in-house advocates to sell the concept, lead the implementation, and provide a critical mass of management approval.

The federal government has been a major customer for existing CAD/CAM products and for highly specialized, one-of-a-kind systems. This latter role has catapulted key government agencies (notably the U.S. Department of Defense and the National Aeronautics and Space Administration) into areas of systems design and management, so that they have become catalysts for advances in knowledge through acquisition, internal research, and sponsored research.

Clearly, our view of these problems is at a point of transition. During the past decade agencies of the federal government have embarked on several programs that have begun to focus attention on both component development needs (the National Bureau of Standards and the U.S. Department of Defense) and system integration problems (the U.S. Air Force and the National Aeronautics and Space Administration). An important factor in these programs is the use of consortia of industrial companies and universities, although university involvement has been limited.

There are a few recent examples of university/industry collaboration at schools such as Rensselaer Polytechnic Institute, Carnegie-Mellon University, and Stanford University. Other universities are attempting to develop similar collaborative efforts, many of them in local areas. It appears that an effective way to conduct research and transfer technology in certain areas would be via these consortia.

Over the years, individual companies have supported applied and generic research at universities through grants of money and equipment. This activity has now been intensified, the most recent development being IBM's announcement in July 1983 of $10 million to five universities for manufacturing-related curriculum development. This was preceded by a $40 million equipment grant to 20 universities.

Individual companies within industries are beginning to explore the feasibility of creating joint R&D partnerships patterned after Microelectronics and Computer Technology Corporation, which was created within the electronics industry to undertake the study of generic research issues.

The central issue in the field of computers in design and manufacturing is a pervasive lack of scientific knowledge. This lack of knowledge is less pervasive than it was 10 years ago, in part because university research has been able to make some significant accomplishments with very limited resources. However, major gaps still exist, as this report illustrates. The first priority of this field is to build a genuine, classical research community. The large integration efforts identified in this report will require a much stronger base of fundamental knowledge than is now available.

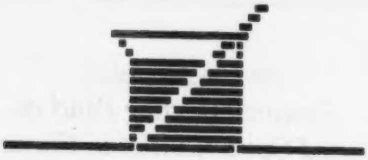
EDUCATION AND TECHNOLOGY TRANSFER

As important as the development of technology is the transfer of that technology to an educated work force (including engineers, managers, and operators). There are four basic issues, each requiring a different mechanism. The first issue is the transfer of prototype technology from R&D laboratories to industrial use. University-industry consortia can facilitate this transfer, as will the traditional mechanisms of conferences, publications and product sales.

The second problem is to educate a new breed of engineers who thoroughly understand all aspects of computer-integrated manufacturing (CIM) engineering in its broadest sense. This involves developing new courses and programs and implies a knowledgeable faculty.

The third issue is the integration of modern computing tools into the traditional program of engineering education. These students will become the users (not developers) of CIM.

The last issue is the training of operators and technicians in an effort to upgrade the existing work force. It is necessary to find cost-effective ways of training people at a significantly faster rate and to develop a means of transferring experience without incurring the high cost of one-on-one instruction. Better productivity in training for skill development is a critical need.



Report of the Research Briefing Panel on Selected Opportunities in Chemistry

Research Briefing Panel on Selected Opportunities in Chemistry

George C. Pimentel (*Chairman*), University of California, Berkeley
Allen J. Bard, University of Texas
John H. Birely, Los Alamos National Laboratory
Michel Boudart, Stanford University
Ronald Breslow, Columbia University
Harry B. Gray, California Institute of Technology
Ralph Hirschmann, Merck Sharp & Dohme Research Laboratories
Alan Schriesheim, Exxon Research & Engineering Company
Howard E. Simmons, E. I. du Pont de Nemours & Company
Barry M. Trost, University of Wisconsin
George M. Whitesides, Harvard University

In cooperation with the NRC Committee to Survey the Chemical Sciences and the Board on Chemical Sciences and Technology

Staff

William Spindel, *Executive Secretary*, Board on Chemical Science and Technology (BCST)
Mary E. Bundy, *Administrative Associate*, BCST
Peggy J. Posey, *Staff Associate*, BCST
Robert M. Simon, *Rapporteur*
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

Table of Contents

Executive Summary 69

I. Introduction 73

II. Priority Opportunities 74
  A. Understanding Chemical Reactivity, 74
    Molecular Dynamics, 75
    Reaction Pathways, 76
  B. Chemical Catalysis, 79
    Heterogeneous Catalysis, 80
    Homogeneous Catalysis, 81
    Photocatalysis and Electrocatalysis, 83
    Artificial-Enzyme Catalysis, 84
  C. Chemistry of Life Processes, 86
    Receptor-Substrate Interactions, 87
    Membranes and Vectorial Chemistry, 87

III. Conclusions 89

Appendix I
  Fundamental Questions Before Us 90

Report of the Research Briefing Panel on Selected Opportunities in Chemistry

EXECUTIVE SUMMARY

This is a time of special opportunity for intellectual advance in chemistry. It derives from our developing ability to probe and understand the elemental steps of chemical change and, at the same time, to deal with extreme molecular complexity. Powerful instrumental techniques are the crucial dimension. They account for the recent acceleration of progress that gives chemistry unusual promise for high return from additional resources.

Chemistry is a central science that provides fundamental understanding needed to deal with most societal needs: to feed the world population, to tap new sources of energy, to clothe and house humankind, to provide renewable substitutes for dwindling or scarce materials, to improve health and conquer disease, to add to our national security, and to monitor and protect our environment. Further, there can be no doubt that chemistry, with its current $12 billion positive balance of trade, is a crucial element in this nation's economic well-being. Our future international competitiveness will depend upon maintaining our present leadership position in the chemical sciences.

Directions of particular promise are becoming apparent in the course of the NAS/NRC study by the Committee to Survey the Chemical Sciences. This Committee is formulating a multi-year program to exploit the richest potentialities in a priority sequence that will maximize return on the needed additional federal investment. We identify here three areas, those that will yield the greatest rewards and furnish the appropriate beginning to this program.

A. Understanding Chemical Reactivity
B. Chemical Catalysis
C. Chemistry of Life Processes

A. *Understanding Chemical Reactivity*

We propose an initiative to apply the full power of modern instrumental techniques and of chemical theory to the clarification of factors that control the rates of reaction and to the development of new pathways for chemical change.

- *Molecular Dynamics*: to elucidate the entire course of chemical reactions, including the unstable atomic arrangements intervening between reactants and products.

- *Reaction Pathways*: to accelerate our rate of discovery of new reaction paths and new classes of compounds.

A principal objective of this initiative will be to provide the fundamental basis for U.S. leadership in development of new processes, new substances, and new materials.

B. *Chemical Catalysis*

We propose an initiative to apply the techniques of chemistry to obtain a molecular-level understanding of catalysts. Ultimately, new catalyst systems will result that will lay the foundation for the development of new chemical technologies. The program will emphasize research in four areas:

- *Heterogeneous Catalysis*: to apply the powerful new instrumental techniques of surface science to the study of chemistry on the surface of a solid.
- *Homogeneous Catalysis*: to take advantage of recent developments in synthetic chemistry that show promise as new soluble catalysts.
- *Photocatalysis and Electrocatalysis*: to investigate solution chemistry catalytically assisted by electrode processes, with and without absorption of light.
- *Artificial-Enzyme Catalysis*: to couple the chemists' ability to control molecular topography with the biochemists' understandings of natural catalysts so as to generate a new class of artificial enzyme-like catalysts tailored to specific needs.

A principal objective of this initiative will be to provide the fundamental knowledge and creative manpower required for the U.S. to maintain competitive advantage in catalysis-aided technologies.

C. *Chemistry of Life Processes*

We propose an initiative to develop and apply the techniques of chemistry to the solution of molecular-level problems in life processes. Research at this border between chemistry and biology requires individuals broadly competent in both areas, and a special effort must be made to develop such individuals. The program will emphasize the molecular aspects of six areas:

- Enzymology (understanding the molecular interactions responsible for enzymatic activity; production of natural enzymes for use as catalysts in chemical synthesis).
- Immunochemistry (the chemical basis of the immune and allergenic response; use, function, and modification of monoclone antibodies; synthesis of antigens and adjuvants).
- Chemical Endocrinology (synthesis of hormones and hormone analogs, especially those related to polypeptides).
- Neurochemistry (molecular basis of nerve transmission; neurotransmitters; agonist and antagonist chemistry; membrane polarization).
- Membrane Chemistry and Vectorial Chemistry (bioenergetics; active and passive transport).
- Biological Model Studies (host/guest chemistry; semisynthetic enzymes; properties of aqueous solutions; active site modeling).

A principal objective of this initiative will be to accelerate the conversion of qualitative biological information into techniques and substances useful in biotechnologies, in human and animal medicine, and in agriculture.

In the next two decades there will be dramatic changes in our basic understandings of chemical change and our ability to marshal those understandings to deliberate purpose. The program presented here is intended to define a leadership role for the U.S., as these advances are won. The rewards accompanying such leadership are commensurate with the prominent role of chemistry in addressing societal needs, in ameliorating the problems of our technological age, and in sustaining our economic well-being. The costs of falling behind are simply intolerable.

The need for enhanced federal investment in chemical research is rooted in a pattern of funding historically appropriate to a test-tube and bunsen burner era, an era long since eclipsed. The sophistication of a modern chemistry laboratory requires a much more vigorous financial commitment, both in capital investment and in supporting services. Even so, the cost is miniscule compared to the stakes involved; we must nurture a $175 billion industry that maintains a $12 billion positive balance of trade. We must provide it with a full reservoir of fundamental knowledge and we must attract a substantial share of our brightest young scientists to this rewarding task. This program can accomplish these goals.

Report of the Research Briefing Panel on Selected Opportunities in Chemistry

I. INTRODUCTION

This is a time of special opportunity for intellectual advance in chemistry. It derives from our developing ability to probe the elemental steps of chemical change and, at the same time, to deal with extreme molecular complexity. Powerful new instrumental techniques are the crucial dimension. They account for the recent acceleration of progress that gives chemistry unusual promise for high return from additional resources.

Chemistry is a central science that provides fundamental understanding needed to deal with most societal needs. It is a critical component in Man's attempt to feed the world population, to tap new sources of energy, to clothe and house humankind, to provide renewable substitutes for dwindling or scarce materials, to improve health and conquer disease, to add to our national security, and to monitor and protect our environment. Basic research in chemistry will surely help future generations cope with their evolving needs and unanticipated problems.

Chemistry is also a crucial element in this nation's economic well-being. The U.S. Chemical and Allied Products industry employs over a million people, makes manufacturing shipments totaling about $175 billion, and currently displays a $12 billion positive international balance of trade, second highest of all commodity groups. Our competitiveness in a range of international markets depends upon maintaining our present position of leadership in the chemical sciences. There is no area of basic science that offers a more secure investment in the Nation's future.

Thus, rapid advances are possible now in chemistry, advances that will continue to enrich our cultural heritage, that will ultimately respond to human needs, and that will sustain our economic competitiveness. Directions of particular promise have become apparent in the course of the NAS/NRC study by the Committee to Survey the Chemical Sciences. This Committee is formulating a multi-year program to exploit the potentialities in a priority sequence that will maximize return on the needed additional federal investment. We identify here three areas that furnish the appropriate beginning to this program.

A. Understanding Chemical Reactivity
B. Chemical Catalysis
C. Chemistry of Life Processes

II. PRIORITY OPPORTUNITIES

Three priority thrusts are described. The first is connected with opportunities to understand, in the most fundamental sense, chemical reactivity and to control it. The second is to advance our understanding of catalysis in all of its facets. The third will extend to the molecular level our understanding of life processes. These three thrusts are complementary and reinforcing in the sequence given. They have different rationales and distinct budget needs, so they are presented separately.

A. *Understanding Chemical Reactivity*

1. *Program.*

We propose an initiative to apply the full power of modern instrumental techniques (lasers, molecular beams, computers, etc.) and of chemical theory to the clarification of factors that control the course of a chemical reaction and to the development of new pathways for chemical change.

- *Molecular Dynamics*: to elucidate the entire course of chemical reactions, including the unstable atomic arrangements intervening between reactants and products.
- *Reaction Pathways*: to accelerate our rate of discovery of new reaction paths and new classes of compounds.

2. *Objectives.*

- To advance our fundamental understanding of chemical reactivity.
- To attract exceptional young scientists to the frontiers of this understanding.
- To apply this understanding to the development of new types of reaction and new classes of compounds.
- To provide the fundamental basis for U.S. leadership in development of new processes, new substances, and new materials.

3. *Rationale.*

Public support for basic research is justified, in part, on cultural grounds and, in part, on practical grounds. Our culture includes the premise that learning about ourselves and our environment is an ample basis for encouragement of scientific investigations. In addition, we recognize that our increasing ability to meet the common desire for survival, comfort, health, and freedom from toil is rooted in our increasing comprehension of what goes on within and around us.

Within the cultural basis for science, three prime questions stand out: the structure of the Universe, the makeup of matter, and the mechanism of life processes. The third is probably the strongest human preoccupation: the nature and preservation of life. Since all life processes—birth, growth, reproduction, aging mutation, death—are manifestations of chemical change, understanding chemical reactivity can be seen as the foundation for our ultimate understanding of life. Thus chemistry, along with biology, can properly be placed with astronomy and high energy physics as scientific investigations justifiable because they contribute to human knowledge in areas of universal philosophical significance.

However, that does not complete the justification for increasing our understanding of chemical reactivity. It is just as convincingly based upon practical grounds. The development of new processes and new materials depends upon our understanding and control of chemical change. Fundamental advances here will present avenues for innovation that can be exploited in the study of combustion, corrosion and electrochemistry, polymer formation, tailored organic molecules, and new solid state materials. It responds to our need to develop new energy sources, to provide alternate materials for nonrenewable or scarce materials, and to maintain economic competitiveness, all while attentively protecting our environment. It permits the U.S. to produce billions of pounds of organic chemicals, at low cost, in high yield, and with minimum waste product. Combined with our 1981 production of 9.8 billion pounds of synthetic fibers (such as polyesters), 28 billion pounds of plastics (such as polyethylene) and 4.4 billion pounds of synthetic rubber, this capability accounts for the total business volume mentioned earlier, an amount exceeding $175 billion with its current international trade surplus of $12 billion.

This is surely a time of special opportunity to deepen our fundamental knowledge of why and how chemical changes take place. The advance of the frontiers of molecular dynamics has undergone a revolutionary acceleration during the last decade. At the same time, synthetic chemists are exponentially increasing the number of known molecular structures, and constantly adding to our arsenal of reaction types and classes of compounds. Much of this can be attributed to the development and application, during the last two decades, of an array of powerful instrumental and analytical techniques. A few examples show that we have new capabilities in hand with which to probe far beyond current frontiers.

Molecular Dynamics

Lasers by themselves have spectacularly expanded our experimental horizons over the last decade. Their short pulse durations permit easy probing of chemical reactions that occur in less than a millionth of a second down to times a thousand-fold shorter (from a microsecond to a nanosecond). With more complex (and more expensive) equipment, chemists are now entering a time domain still another thousand-fold shorter (to a picosecond). Lasers also provide tunable, extremely narrow frequency light sources, implying vastly greater diagnostic sensitivity and selectivity. Finally, lasers give us extremely high light intensities, either in the form of short pulses of ultrahigh power or as continuous light sources with unprecedented photon flux.

Computers also make important contributions to this revolution, both in experimental and theoretical chemistry. They are now common components of experimental assemblies and a number of the chemist's more sophisticated instruments have built-in (dedicated) computer capability as an essential element. (For example, fourier transform infrared spectrometers, fourier transform nuclear magnetic resonance spectrometers, and X-ray spectrometers). Furthermore, today's computing capability has brought *ab initio* calculations and chemical theory to new levels of fruition.

These are only exemplars—a host of sophisticated instrumental methods ought to be mentioned: molecular beams, ion cyclotron resonance, electron spin resonance, photoelectron spectroscopy, magnetic circular dichroism, resonance Raman, Mossbauer, etc. All of these account for the rich possibilities lying before us. However, the historical basis for funding basic research in chemistry was not premised on the existence (and cost) of such sophisticated equipment. One of the strongest arguments for a large increase in federal support for chemical research is that existing levels do not provide sufficient access to these state-of-the-art techniques and without them, U.S. science will fall behind in areas crucial to our national well-being.

Turning to these new frontiers, the advance of understanding of *molecular dynamics* is one of the most rapidly moving. Here are some of the new phenomena we can now explore.

- *Fast chemical processes in real time.* The primary processes of numerous chemical changes are determined by dynamical events that are intrinsically fast. Many solution phase reactions, photochemical reactions, electron solvation, molecular reorientation in liquids, and reactions of certain biological systems can now be watched "in real time" on the nanosecond-to-picosecond time scale.
- *Energy transfer and movement.* In all chemical changes, the pathways for energy movement are determining factors. Competition among these pathways (including energy dissipation) determines the product yields, the product-state distributions, and the rate at which reaction proceeds. Chemists are

now able to track these pathways: radiation, internal conversion, intersystem crossing, intramolecular and intermolecular energy transfer, unimolecular decomposition.

- *Ab initio calculations of reaction surfaces.* With today's computers, the structure and stability of any molecular compound with up to 3 first-row atoms (carbon, nitrogen, oxygen, fluorine) plus various numbers of hydrogen atoms can be calculated almost to the best accuracy available through experiment. This opens to the chemist many situations not readily accessible to experimental measurement. Short-lived reaction intermediates, excited states, and even saddle points of reaction can now be understood, at least for small polyatomic molecules. In a major advance, we can now calculate the forces on all of the atoms during their reorganization from reactant to product molecule geometries.
- *State-to-state chemistry.* Major developments in molecular beam technology, including mass spectrometric ("universal") detectors, vacuum technology, laser excitation, and supersonic jet sources, have enormously broadened their applicability. In the ultimate experiment, reactants can be brought together in precisely known energetic states, with every quantum number fixed, and then the product energies and relative probabilities can be established in the same detail.
- *Mode-selective chemistry.* With high-power, sharply tunable lasers, it is possible to excite one particular degree of freedom for many molecules in a bulk sample. As long as this situation persists, such molecules react as if this degree of freedom is at a very high temperature while all the rest of the molecular degrees of freedom are cold. The chemistry of such selective excitation could potentially reveal the importance and role of that particular degree of freedom in facilitating reaction. To extract this valuable information, energy redistribution and relaxation must be brought under control, and this frontier is being explored.
- *Multiphoton excitation.* One of the most surprising discoveries of the 1970's was that powerful pulsed lasers could highly excite the vibrational degrees of freedom of a molecule on a time scale short compared to molecular collision times. With many tens of quanta of vibrational excitation, the molecule displays unusual chemical pathways. This opens a new type of photochemistry with unique possibilities, the first of which to be recognized was isotope enrichment.

The rich potentialities of these techniques promise to show us how reactions proceed over their entire course. We will at last be able to understand the unstable situations intervening between reactants and products. To realize these benefits, there must be wider access to the sophisticated instrumentation (including computing capability) and the cadre of bright young scientists attracted to this area must be encouraged and increased.

REACTION PATHWAYS

A manifestation of this increasing understanding and control of chemical reactivity is the rapid advance now taking place in devising new reaction pathways in synthetic chemistry. This progress presents a high leverage opportunity since herein lies the foundation of future developments of new products and new processes.

Again powerful instrumental techniques play a central role. The rapid and definitive identification of reaction products, both in composition and structure, accounts for the speed with which synthetic chemists are able to test and develop their synthetic strategies. The nuclear magnetic resonance and mass spectra show what elements are present and the structural environment of virtually every atom. The X-ray crystal structure reveals the complete molecular structure: the interatomic distances, bond angles, and even the right- or left-handedness when mirror-image relationships are present. Spectroscopic techniques, as well, have been essential to progress in the rapidly developing area of organometallic chemistry. The visible and infrared spectra of transition metal complexes reveal electronic configurations and bonding, the foundation for clarifying mechanisms for ligand substitution and electron-transfer processes.

- *Selective pathways in organic synthesis.* Selectivity is the key challenge to the synthetic chemist—to make a precise structural change to a single desired product. The different intrinsic reactivity in each bond type must be recognized (chemoselectivity), reactants must be brought together in proper orientation (regioselectivity) and the desired three-dimensional spatial relations must be obtained (stereoselectivity). The degree to which this type of control can be achieved is shown in the synthesis of the substance adamantane ($C_{10}H_{14}$). This novel molecule resembles in structure a 10-atom "chip" off of a diamond crystal. In a *tour de force*, it was originally produced by a many-step process in only 2.4% yield. Recent research in polycyclic hydrocarbon synthesis now gives adamantane in one step in 75% yield. Then, a surprise practical payoff came when it was discovered that adding a single amine substituent to adamantane gave adamantine (1-amino-adamantane), which proved to be an antiviral agent, a prophylactic drug for influenza, and a drug to combat Parkinson's disease.

The recent development of new ways to construct cyclic five-membered rings has again paid off handsomely. The diverse array of applications includes production of thienamycin, $C_{13}H_{14}N_2SO_4$, a novel relative of penicillin and an important new drug. At another extreme, large ring compounds have been exceptionally difficult to synthesize. Their structures are complicated by functionally crucial left/right-handed structural geometrics ("chiral centers"). Their wide ranging biological properties—from pleasant fragrances for perfumes to anti-fungal, antitumor, and antibiotic activities—make large ring synthesis an interesting challenge. An example is erythromycin, $C_{37}H_{68}O_{12}N$, which in the desired atomic hookup can be shaped into 262,144 different structures derived from the many possible ways to couple the right- and left-handedness at its 18 chiral centers ($2^{18} = 262,144$). Twenty-five years ago, this compound was judged to be "hopelessly complex" by R. B. Woodward, who won the Nobel Prize for synthesizing molecules as complex as quinine and vitamin B_{12}. Today we can aspire to such a goal, in part because of the development of specially designed templates that bring together the terminal atoms of a fourteen-atom chain to form a fourteen-membered ring. This provides the structural framework of erythromycin and it has already resulted in the synthesis of a number of constituents of musk and contributed to our understanding of smell.

- *Crossing inorganic/organic boundaries.* The traditional line of demarkation between organic and inorganic chemists has virtually disappeared as the list of fascinating metal-organic compounds continues to grow. The ubiquitous appearance of these compounds in biological systems underscores the importance of encouraging this cross-boundary research. Furthermore, research in developing new inorganic substances has provided a surprising dividend in their frequent applicability in organic synthesis.

The latter situation is illustrated by the borohydrides. The cohesive picture we have, at last, for this strange boron/hydrogen family was not possible before their study by X-ray crystallography, infrared and NMR spectroscopy, and molecular orbital theory. Now borohydrides are widely used as selective, mild reducing agents in organic synthesis. Silicon and transition-metal organometallic compounds give other examples. Silicon compounds, for example, are used to fold an extended molecular reactant precisely as needed to synthesize the molecule cortisone. Now this valuable therapeutic agent can be made in less than 20 steps at a yield 1000 times higher than achieved in the earlier 50-step process.

Compounds with metal atoms sandwiched between organic rings continue to proliferate and their chemistry will likely relate to catalysis. The metal-organic complexes provide another active synthetic field, with interest both in catalysis and, perhaps, in radioactive or metal-waste cleanup.

- *Novel solids.* Chemists are learning how to prepare solids with quite remarkable properties. Thus, families of solid substances are being synthesized which, like the alumina-silicates (the "zeolites"), are deliberately contrived to include particular shaped cavities and channels. Guest molecules that slip comfortably into these channels can be held in favored conformations as reactants bring about desired chemical changes. The result is that only particular molecules (the ones that fit) react and they do so in a structurally specific way. Entirely different sorts of solids are giving new families of electrically conducting materials. The lead molybdenum sulfide, $PbMo_6S_8$, is a superconductor that can remain superconducting up to 600 kilogauss fields. One of the several conducting organic materials is polyacetylene which, on exposure to various chemicals (iodine, arsenic trifluoride, etc.), can approach metallic conductivities. Polysulfurnitride, $(SN)_x$, not only rivals metal conductivity, it also displays superconductivity. Another promising class of new linear-chain polymer conductors are those based upon metallo-macrocycle complexes held together by bridging groups and treated with an electron-accepting compound. This class is amenable to a wide range of electronically significant chemical modification, hence tailoring to intended application.

Pressure as a controlling dimension of solid state reactions has long been recognized but little understood. With techniques now available, solids can be compressed and studied spectroscopically up to a few hundred thousand atmospheres pressure. New molecular structures, reactivities and electronic properties can be obtained. Photochromic materials provide an example. Substances like the anils, spiropyranes, and bianthrones change color when exposed to light. Under pressure, the colors can be "tuned" and some of these materials will display the color change merely by heating, without the need for exposure to light. These effects are attributable to changes in molecular structures and electronic makeup which, in turn, can affect chemical reactivity.

Solid state chemistry is relatively less active in the U.S. than abroad; French, German, and Japanese chemists have leadership positions. Hence, there can be a high return if more activity can be stimulated here in the synthesis and characterization of radically new solids. There will emerge new semiconductors, solid state ionic materials (used in batteries, memory devices, display devices, and chemical sensors), ferroelectrics, pyroelectrics, and magnetic materials. Creative synthetic chemists must be attracted and encouraged in this area.

- *Pathways using light as a reagent.* Another promising chemical pathway is connected with the use of light photons in chemical synthesis. Many organic molecules display quite different chemistry and structures after absorption of light. Photosynthesis provides the most striking prototype and our understanding is advancing rapidly. "Artificial photosynthesis" both mimicking Nature and generalizing into new directions is under active study. Chemical storage of solar energy is an obvious long-term goal. In addition, new routes in chemical synthesis are offered. Some high energy ("strained") molecular structures, including those of many natural products, must be formed in energy-consuming (endothermic) processes (e.g., the mycin antibiotics, certain alkaloids, vitamin D precursors, and certain steroid hormones). It is difficult to put in this energy because the active reagents tend to threaten the fragile product. In the photochemically induced process, less aggressive reactants can be used which can be excited with light to approach the product gently from above, with delicate control through the photolysis wavelength. A single example shows the great potential. A calcium homeostatic steroid hormone with the formula $C_{26}H_{44}O_2$ (1,25-dihydrovitamin D_3) can be synthesized photochemically. By using tuned laser light controlled to the optimum wavelength (300nm), the synthetic yield of the desired structure has been increased 400%. How important such light-assisted chemistry will be remains to be seen. Chemists understand the ground state (unexcited) chemistry of literally millions of compounds. Virtually every one of these compounds has different chemistry after light absorption but we have explored only a miniscule fraction of this new domain!

4. *Collaboration*

The thrust toward *understanding chemical reactivity* has inviting avenues for collaboration across traditional disciplinary boundaries. Much of the most fundmental work on molecular dynamics is at the chemistry-physics interface and it specially benefits from close interaction between chemists and physicists on the one hand and between experimentalists and theoreticians on the other. In a similar way, exciting new horizons in synthesis are bringing together the perspectives and techniques of individuals who used to be differentiated as inorganic or organic chemists. Collaborative mechanisms should be encouraged.

B. *Chemical Catalysis*

1. *Program.*

We propose an initiative to apply the techniques of chemistry to obtain a molecular-level understanding of catalysts. Ultimately, new catalyst systems will result that will lay the foundation for the development of new chemical technologies. The program will emphasize research in four areas:

Heterogeneous catalysis: to apply the powerful new instrumental techniques of surface science to the study of chemistry on the surface of a solid.

- *Homogeneous catalysis*: to take advantage of recent developments in synthetic chemistry that show promise as new soluble catalysts.
- *Photocatalysis and electrocatalysis*: to investigate the rich possibilities of solution chemistry catalytically assisted by electrode processes, with and without absorption of light.
- *Artificial-enzyme catalysis*: to bring together chemists' ability to synthesize molecules of predesigned topography and the biochemists' emerging understandings of natural catalysts, the enzymes. We should be able to generate a new class of artificial enzyme-like catalysts tailored to specific needs.

2. *Objectives.*

- To identify the fundamental molecular ingredients of catalytic processes in order, ultimately, to guide the development of chemical high technology.
- To synthesize new catalysts that may serve as the foundation of new chemical processes.
- To increase the number of young scientists who have the breadth and background needed to contribute to the field of chemical catalysis.
- To provide the fundamental knowledge and creative manpower required for the U.S. to maintain competitive advantage in catalysis-aided technologies.

3. *Rationale.*

A catalyst accelerates chemical reactions toward equilibrium without being consumed. This acceleration can be as much as ten orders of magnitude. A *selective* catalyst can have this same dramatic effect but on only one of many competing reactions. While the development of new catalysts was empirical fifteen years ago, research innovations in chemical sciences over the last ten years are converting catalysis from art to science. A familiar example displays the benefits to be gained—the catalytic converter developed for automobiles to reduce air pollution. The catalyst—containing two grams of platinum, palladium, and rhodium per car—oxidizes unburned hydrocarbons and carbon monoxide to harmless water and carbon dioxide. Simultaneously, it reduces toxic nitrogen oxides to harmless nitrogen gas.

Estimates that twenty percent of the gross national product is generated through the use of catalysis demonstrate the major role industrial catalysis has in satisfying such diverse societal needs as food production, energy conversion, defense technologies, environmental protection and health care. On the horizon is the extensive use of catalysts to tap new energy sources—such as heavy oils, coal, oil shale, tar sands, lignites, and biomass—as we run out of oil. However, American preeminence in catalysis science is now challenged by Japan, Germany, the USSR, and France. These countries have established institutional networks to accelerate research and development in this evolving high-technology field. To meet this challenge, we must keep our own catalytic science vital.

Industrial research in catalysis is critically dependent on the techniques and concepts which have been developed in university and National Laboratories. Since industrial laboratories use the latest surface science and laser spectroscopic techniques, they want their new young scientists to bring with them state-of-the-art experience. Furthermore, industrial catalysis research must be chosen with an eye toward existing technology and product lines. Industry cannot sustain the adventurous fundamental research that will underlie the discovery of radically new catalysts. To illustrate this, consider the enormously important conversion of atmospheric nitrogen into ammonia. The original work by Haber in Germany produced an ammonia yield of 13% per pass. Today the same process is used on a massive scale. Yet, seventy years after Haber's work, a modern ammonia unit producing some 1500 tons per day operates at only slightly greater yield per pass (perhaps 15%). There is no physical limitation here—equilibrium is on our side. What is needed is fundamental research in catalysis. This is a role for university chemists. It is a place for federal investment.

This chemical catalysis program will concentrate on long-term research to learn about catalytic processes on the molecular level. Then the new fundamental knowledge generated will be transferred into technology by the employment of the emerging young scientists in industry and by the ties formed in this program among scientists in the universities, the National Laboratories and industry.

HETEROGENEOUS CATALYSIS

Heterogeneous catalysts are solid materials prepared with large surface areas (1-500 m^2/g) upon which chemical reactions occur at extremely high rate and selectivity. Some major new commercial processes based on heterogeneous catalyst developments in recent years include:

- "reforming" hydrocarbons to high octane composition (platinum alloy catalysts);
- "cracking" high paraffins to gasoline (molecular sieve catalysts);
- methanol to gasoline, jet fuel (molecular sieve catalysts);
- exhaust converter on automobiles (platinum, palladium, rhodium catalysts);
- ethylene oxidation to ethylene oxide (silver, cesium, chlorine catalyst);
- propylene oxidation to acrolein and acrylonitrile (bismuth, and molybdenum oxide catalyst);
- ethylene polymerization (chromium catalyst);
- propylene polymerization (titanium, magnesium oxide catalyst).

Surface science is developing rapidly and now gives us experimental access to this two-dimensional reaction domain. Because of the unsatisfied bonding capability of the atoms at the surface, chemistry here is dramatically different from that of the same reactants brought together in solution or the gas phase. But when chemists are able to "see" what molecular structures are on the surface, then all of our knowledge of reactions in conventional settings becomes applicable. This will open the door to understanding and controlling the chemistry in this surface domain. There are five areas of heterogeneous catalysis where this understanding will have major impact on new chemical technologies.

- *Molecular sieve synthesis and catalysis.* Molecular sieves are crystalline alumina-silicates containing pores or channels within which chemical reactions can be initiated. They offer unparalleled efficiency both for cracking of petroleum and for conversions such as shale oil or methanol to gasoline. We need to know better how to synthesize molecular sieves with controlled molecular pore size, as well as to determine the elementary reaction steps and intermediates that account for their efficacy.
- *Metal catalysis.* Finely dispersed transition metals are already coming into use to catalyze hydrocarbon conversions and ammonia synthesis for fertilizers. Other such applications and improved performance will follow from intensive research into the control of surface structures, oxidation states, residence time of reaction intermediates, and resistance to catalyst "poisons" (such as sulfur).
- *Substitutes for precious metal catalysts.* Many of the most effective catalysts are rare metals not available in the U.S., including rhodium, platinum, palladium, and ruthenium. Their strategic value requires a concerted research effort to find more accessible substitutes, such as transition metal oxides, carbides, sulfides, and nitrides.
- *Conversion catalysts.* We must find catalysts to convert abundant substances to useful fuels and industrial feedstocks. These reactions include conversion of atmospheric nitrogen to nitrates, methane to methanol, carbon dioxide to formate, and depolymerization of coal and biomass to useful hydrocarbons.
- *Catalysts to improve air and water quality.* We have many environmental pollution problems for which we need to match the spectacular success of the catalytic converter used to clean automobile exhaust gases. To begin, we need catalysts that remove sulfur oxides from smoke plumes, that purify water, and that prevent acid rain.

As we learn more about the molecular structures at the solid-gas interface (reactants, intermediates and products), a better understanding of surface chemical bonding will follow. We can look forward to understanding additives that modify catalyst performance ("promoters" and "poisons"). Then, the challenging subject of synthesis of the designed catalyst can be addressed. All of this fundamental knowledge will underlie and facilitate the development of new and more selective heterogeneous catalysts.

HOMOGENEOUS CATALYSIS

Homogeneous catalysts are soluble and active in a liquid reaction medium. Often they are complex metal-containing molecules whose structures can be modified to tune reactivity in desired directions to achieve very high selectivities. (In this sense homogeneous catalysts can be superior to heterogeneous ones.) The largest industrial-scale process using homogeneous catalysis is the partial oxidation of para-xylene to terephthalic acid (United States production in 1981 was 6.2 billion pounds per year), which uses dissolved salts of cobalt and manganese as the catalyst system. Most of the product ends up polymerized to give us polyester clothing, tire cord, soft drink bottles, and a host of other useful articles.

An important branch of homogeneous catalysis has developed from research in organometallic chemistry. An example is rhodium dicarbonyl diiodide, employed in the commercial production of acetic acid from methanol and carbon monoxide. With this catalyst present, the reaction economically gives more than 99% selectivity to acetic acid.

There are three areas of homogeneous catalysis with potential for major impact on new chemical technologies from increased understanding.

- *Activation of inert molecules.* There are a number of relatively inert molecules that are enticing as reaction feedstocks because of their abundance: nitrogen, carbon monoxide, carbon dioxide, and methane. One way to approach this end might be through homogeneous organometallic catalysis. Dramatic examples of this promise are beginning to appear. Soluble compounds of tungsten and molybdenum with molecular nitrogen have been prepared and induced to produce ammonia under mild conditions. The carbon-hydrogen bonds in normally unreactive hydrocarbons have been split by organorhodium and organoiridium complexes. Hope for build-up of complex molecules from one-carbon molecules, such as carbon monoxide, is encouraged by recent demonstrations of carbon-carbon bond formation at metal centers bound in soluble metal-organic molecules. Synthesis of compounds with multiple bonds between carbon and metal atoms has had a major impact in clarifying the catalytic interconversion of olefins. While there is much to learn, the stakes are high and the odds for success are excellent.
- *Metal cluster chemistry.* An adventurous frontier of catalysis lies in the expanding capability of chemists to synthesize molecules built around several metal atoms in proximity (a "metal cluster"). In parallel, solution and cryogenic techniques are revealing the structures and chemistry of small aggregates containing only metal ions or atoms ("naked clusters"). All of these clusters, bound or "naked", furnish a natural bridge between homogeneous catalysis and bulk metal, heterogeneous catalysis.

  Many directions are being explored. Cubical units of four metal atoms and four sulfur atoms are now known for iron, nickel, tungsten, zinc, cobalt, manganese, and chromium. This "cubane" structure for iron has been found to be the functional unit in the ferrodoxin iron proteins that catalyze electron transfer reactions in biological systems. Many cluster compounds have been made from metals bound to carbon monoxide. These metal carbonyls have formulas $M_x(CO)_y$, and x can be made very large. (The world's record as of this writing is a platinum compound with $x = 38$.) It is intriguing that many of the most catalytically active metals also form cluster compounds (e.g., rhodium, platinum, osmium, ruthenium, iridium, etc.). Now the chemistry of these elements can be studied as a function of cluster size.
- *Stereoselective catalysts.* Another exciting frontier involves the development of homogeneous stereoselective catalysts. Many biological molecules can have either of two geometric structures connected by mirror-image (chiral) relationships, and generally only one of these structures is functionally useful in the biological system. If a complex molecule has seven such chiral carbon atoms and a synthetic process produces all of the mirror-image structures in equal amounts, there would be $2^7 = 128$ structures, 127 of which might have no activity or, worse, some undesired effect. Thus the ability to synthesize preferentially the desired structure with the desired geometry at every chiral center is essential.

  Immense strides in this area are being made. For example, L-dopa, an amino acid that has revolutionized the treatment of Parkinson's disease, is now made using an asymmetric addition of hydrogen to a carbon-carbon double bond. The catalyst is a soluble rhodium phosphine catalyst that gives 96% of the correct structure with high efficiency. Stereospecific oxidations can also be carried out. The recent invention of a titanium catalyst to add, in a specific geometry, an oxygen atom across a carbon-carbon double bond has lowered the price of gypsy moth attractant ten-fold. Despite these successes, the basic factors that produce stereochemical control are not at all well understood. Mechanistic studies are needed and the rewards from better understanding will be great.

Future advances in homogeneous catalysis are dependent upon easy access to advanced instrumental techniques: X-rays, high-field nuclear magnetic resonance, electron spin resonance, mass spectroscopy, and computational facilities. Further, work at the interfaces between organic, inorganic and physical chemistry is involved, so breadth of knowledge is especially important.

Photocatalysis and Electrocatalysis

Exciting advances have recently been made in the study and control of chemistry taking place at the interface between a liquid solution and an electrochemical electrode surface. In some applications, the chemistry is initiated by absorption of light by a semiconductor used as an electrode. Whether light is involved (photocatalysis) or not (electrocatalysis), this rapidly moving field depends upon our knowledge both of homogeneous catalysis and of semiconductor behavior.

- *Photoelectrochemical cells.* There is considerable promise from research directed toward conversion of light to electrical energy in a photoelectrochemical cell. In such a cell, one or both of the electrodes is a semiconductor material that absorbs the incident light. The lure is that semiconductors, probably with chemically modified surfaces, can absorb red and infrared light, the spectral region within which most solar energy falls.

Chemists are learning how to bond dyes and protective polymers to the surface of a semiconductor solid through covalent bonds. There is now the potential for solving the key problems of photoelectrochemical cells: to catalyze the essential electron-transfer reactions with solution species, to prevent chemical erosion of the electrode, to avoid back-reaction, and to shape the active spectral region to match that of the energy source, the sun. Already cells for conversion of solar energy to electrical energy have rivaled those of solid state photovoltaic cells (e.g., the n-GaAs/selenide system operates at 14% efficiency). Numerous "non-conventional" semiconductors are under investigation, such as TiO_2, CdS, WSe_2. These materials are still speciality products but there is good prospect that they will cost much less than existing single crystal photovoltaic materials.

- *Photocatalysis.* Related studies involving photoelectrochemical cells or suspensions of semiconductor materials focus attention on the chemistry that can be brought about. In these systems, light absorbed in the semiconductor promotes catalytic oxidation-reduction chemistry at the electrode-solution or at a membrane-solution interface. Such oxidation-reduction chemistry has significant scientific interest and, likely, great practical significance. For example, photodestruction of a toxic waste material such as cyanide has been demonstrated at titanium dioxide surfaces. A more popularized and perhaps feasible concept is that such photocatalytic chemistry, solar energy driven, could produce massive amounts of hydrogen and oxygen from the photoelectrolysis of water. What an intriguing prospect, to convert from diminishing and seriously polluting petroleum fuels to a renewable fuel that burns to water and that is made from water using energy from the sun.

- *Electrocatalysis.* Apart from light-initiated processes, electrode surfaces with catalytic activity offer a new domain for chemical synthesis. In a field with a long heritage, recent developments have shown that electrode surfaces can be chemically tailored to promote particular reactions. This research area has adopted techniques from the semiconductor industry (such as chemical vapor deposition) and coupled them with imaginative synthetic chemical techniques for surface modification.

An example is the electrocatalyst family developed for use in chlorine generation in chlor-alkali cells. A successful case is based upon a thin layer of ruthenium dioxide (the catalyst) deposited on a base metal electrode. This electrocatalyst has dramatically changed practice in the chlor-alkali industry (an industry representing billions of dollars in sales) because of its improved energy efficiency and reduced cell maintenance. Future developments will include radically improved fuel cells, which provide clean and thermodynamically efficient conversion of chemical fuels to electricity.

- *Chemistry at the solid/liquid interface*. Before the technological potentialities of any of the above can be realized, we must have a much better understanding of the nature of chemistry at the semiconductor/liquid interface. The marvelous instrumentation developed for surface science studies is, of course, applicable only at solid/vacuum interfaces. We need comparable capability at the solid/liquid boundary. This capability will be won from fundamental research in solid state chemistry, electrochemistry, surface analysis, and surface spectroscopy. Where the biggest gains are to be made is a speculative question. The surprising discovery of the million-fold intensification involved in surface-enhanced Raman effect encourages optimism.

The potential gains from these exciting areas are considerable. We need to know how to catalyze multielectron transfer reactions at an electrode surface. That is the chemistry required, for example, to photogenerate a liquid fuel such as methanol from carbon dioxide and water. Multielectron transfer catalytic electrodes for oxygen reduction in electrochemical cells would find a welcoming home in the fuel cell industry.

It is also likely that research on semiconductor electrode surface modification will reflect back beneficially into the field of electronics. Thus, the integrated circuit technology based upon GaAs may depend upon control of its surface chemistry. Already photoresist/chemical etching techniques are recognizing the importance of the chemistry involved in surface modification (witness the pursuit of "anisotropic chemical etching").

In summary, our evolving understanding of the electrode/solution interface, buttressed by concepts based on semiconductor electrodes and the development of a number of new methods for modifying electrode surfaces, has provided powerful new approaches to both photocatalysis and electrocatalysis. Future advances will benefit synergistically from progress in heterogeneous and homogeneous catalysis, increased understanding of mass and charge transport within the electrode surface layers, and continued development of experimental methods and theoretical models for the interface.

ARTIFICIAL-ENZYME CATALYSIS

The most striking benefit from our bounteous knowledge of reaction pathways and the analytical capacity of modern instrumentation has been the development of capability to deal with molecular systems of extreme complexity. With the synthetic chemist's prowess and such diagnostic instruments as nuclear magnetic resonance, X-ray spectroscopy, and mass spectroscopy we can now synthesize and control the structure of molecules that approach biological complexity. This includes the ability to fix the molecular shape, even extending to the mirror-image properties that are so crucial to biological function.

There is no application of these capabilities more intriguing than coupling them with our growing knowledge of catalysis to synthesize artificial enzymes. In Nature, enzymes are the biological catalysts that accelerate a wide variety of chemical reactions at the modest temperatures at which living organisms can survive. A given enzyme selects from a many-component system a single reactant molecule and transforms it to a single product with prescribed chiral geometry.

Without catalysts, many simple reactions are extremely slow under ambient conditions. Raising the temperature speeds things up, but this may include a variety of undesired outcomes—acceleration of unwanted reactions, destruction of delicate products, and waste of energy. Hence, there are strong reasons to develop synthetic catalysts that work like enzymes. First, natural enzymes do not exist for most of the chemical reactions in which we have interest. In the manufacture of polymers, synthetic fibers, medicinal compounds, and many industrial chemicals, very few of the reactions used could be catalyzed by naturally occurring enzymes. Even where there are natural enzymes, their properties are not ideal for chemical manufacture. Enzymes are proteins, sensitive substances that are easily denatured and destroyed. In those industries that do use enzymes, major effort is directed toward modifying them to make them more stable.

- *Controlled molecular topography and designed catalysts.* We have a pretty good idea how enzymes work. Nature contrives a molecular surface fitted to a specific reactant. This surface attracts from a mixture the unique molecular type desired and immobilizes it. Held on the surface, the selected reactant is forced to abandon its flexibility and assume a distinct shape. When the reaction partner arrives, the scene is set for the desired reaction to take place in the desired geometry.

Organic chemists who have taken up this challenge are making excellent progress. Without special control, large molecules usually have exclusively convex surfaces (ball-like shapes). So a first step has been to learn to synthesize large molecules with concave surfaces. Then, these concave surfaces could be shaped to accommodate a desired reactant. Cyclodextrins provide examples—they are toroidal in shape. The crown ethers, developed over the last 15 years, have a quite different surface topography. For instance, 18-crown-6 consists of twelve carbon atoms and six oxygen atoms evenly spaced in a cyclic arrangement. In the presence of potassium ions, the ether takes up a crown-like structure in which the six oxygen atoms point toward and bind a potassium ion. Lithium and sodium ions are too small and rubidium ions too large to fit in the crown-shaped cavity, so this ether preferentially extracts potassium ions from a mixture. Much more ornate examples now exist. Chiral binaphthyl units can be coupled into cylindrical or egg-shaped cavities. With benzene rings, enforced cavities have been made with the shapes of bowls, pots, saucers, and vases.

Plainly we are moving toward the next step, to build into these shaped cavities a catalytic binding site. The earliest successes will likely be patterned after natural enzymes but there is no doubt that in time, artificial, enzyme-like catalysts will not be limited by what we can find already known in Nature.

- *Biomimetic enzymes.* A short-cut approach is to pattern artificial enzymes closely after natural enzymes. This has been called "biomimetic chemistry." As examples, mimics have been prepared for the enzymes that biologically synthesize amino acids. The artificial enzymes incorporate the important catalytic groups of a natural enzyme (vitamin B_{6}, for example) and show good selectivity, even including the formation of the correct chirality in the product. Mimics have been prepared for several of the common enzymes involved in the digestion of proteins, e.g., substances that catalyze the cleavage of RNA have been synthesized based upon the catalytic groups found in the enzyme ribonuclease. Mimics have also been synthesized which imitate the class of enzymes called cytochromes P-450, which are involved in many biological oxidations, and the oxygen carrier hemoglobin. Furthermore, mimics have been prepared for biological membranes and for those molecules which carry substances through membranes. These have potential applications in the construction of organized systems to perform selective absorption and detection, as in living cells.

It is important for the U.S. to build on its early lead in this field. Although most of the work mentioned above has been done in the U.S., the Japanese have also become extremely active and have specifically targeted "biomimetic chemistry" as an area of special opportunity. Research on synthetic organic chemistry develops novel methods to construct the required molecule, and elaborates new kinds of structures. The study of detailed reaction mechanisms in organic and biological chemistry permits a rational approach to catalyst design. This is an area ripe for development, deserving special encouragement as a part of this program in Chemical Catalysis.

4. *Collaboration.*

Thus, catalysis research with the four facets named here involves many subfields of chemistry (surface chemistry, solid state chemistry, organic synthesis, inorganic synthesis, photochemistry, electrochemistry, bioorganic chemistry and chemical engineering) along with the contiguous fields of solid state physics, surface physics, biology, and biochemistry. Each of these different perspectives is needed to address fully the common goal, to understand how a catalyst intermediary can enhance preferentially the rate at which a desired reaction takes place. Hence, there is great opportunity for collaboration across disciplinary boundaries.

C. *Chemistry of Life Processes*

1. *Program.*

We propose an initiative to develop and apply the techniques of chemistry and the skills of individuals knowledgeable in that discipline to the solution of molecular-level problems in life processes. This initiative includes both the application of chemical science to biology, and the development of new chemical science based on stimuli provided by biology. Research at this border between chemistry and biology requires individuals broadly competent in both areas, and a special effort must be made to develop such individuals.

The program will emphasize the molecular aspects of six areas.

- **Enzymology:** to understand the molecular interactions responsible for natural enzymatic activity and its inhibition; to produce natural enzymes for use in chemical synthesis.
- **Immunochemistry:** to learn the chemical basis of the immune and allergenic response; use, function, and modification of monoclonal antibodies; synthesis of antigens and adjuvants.
- **Chemical Endocrinology:** to synthesize hormones and hormone analogs; to understand hormone agonist and antagonist mechanisms.
- **Neurochemistry:** to determine the molecular basis of nerve transmission, neurotransmitters, agonist and antagonist chemistry, and membrane polarization.
- **Membrane Chemistry and Vectorial Chemistry:** to clarify active and passive transport (bioenergetics).
- **Biological Model Studies:** to understand host/guest chemistry, semisynthetic enzymes, and active sites.

2. *Objectives.*

- To strengthen basic research in the life sciences by providing the molecular-scale analytical, synthetic, and structural techniques required to identify, prepare, and study the organic chemical compounds which are central to modern molecular biology.
- To accelerate the conversion of qualitative biological information into techniques and substances useful in human and animal medicine and agriculture.
- To educate graduate students and postdoctoral fellows in both chemistry and biology; to provide individuals skilled in chemistry (especially organic synthesis, organic reaction mechanisms, and analytical

chemistry), who can understand and interpret biological problems and opportunities in molecular terms.

3. *Rationale*.

The study of the molecular bases of the reactions occurring in life offers enormous opportunity both for discovery in basic physical science and for application to human health, animal health, and agriculture. This area of research includes some of the most important problems which now exist in science. Solutions to these problems will have very large and beneficial impacts on society.

Biology has evolved a range of problems of revolutionary significance that now require analysis in terms of molecular interactions. The molecular synthesis and analysis required for the solutions of these problems can only be provided by chemistry.

A major fraction of biology can be reduced to two broad types of problems at the molecular level:

Receptor-Substrate Interactions

Essentially all biological processes are mediated by the selective interaction of a protein receptor (enzyme, antibody, membrane and intracellular receptors) with one or more specific substrates (enzyme substrate, antigen, hormone, neurotransmitter, or simple molecule or ion).

To study these processes in molecular detail, and to control them with the precision required for rational applications in medicine, agriculture, and other areas, it is necessary to be able to:

- isolate and identify the structures of endogenous substrates or substrate analogs;
- synthesize these substrates and analogs in useful quantities and high purity;
- analyze their interactions with their receptors both in physical-chemical and biological terms;
- modify their structures to give useful agonist and antagonist activity and to facilitate effective delivery into living organisms.

The development of analytical, structural, synthetic, and mechanistic techniques appropriate to this area of molecular science represents an enormous challenge and opportunity for chemistry. The active substances may be present *in vivo* only in minute quantities; they are usually water soluble (polypeptides, polysaccharides) or surface active (lipids, membrane or hydrophobic proteins); they may have high molecular weights and complex or heterogeneous structures. Medicine and biology need the methods of chemistry to manipulate these substances; chemistry cannot presently provide fully adequate methods. Hence, the challenge and the opportunity for chemists.

Membranes and Vectorial Chemistry

The set of chemical reactions that constitute life take place in cells and organelles. The concept of vectorial chemistry (reactions depending upon spatial separation of reactants into regions of different concentrations—that is, on concentration gradients—for their driving force) is poorly understood at a molecular level, but essential to biochemistry (the Mitchell chemiosmotic hypothesis is the best-known example).

Understanding the relations between concentration gradients across membranes and processes which form or break chemical bonds is a formidable task, one which requires active chemical modeling at both the synthetic and the kinetic level to complement the current activity in mechanistic biology. The related problem of the transport of molecules across membranes (for example, between blood and a particular type of cell) is central to the area of pharmacokinetics and drug delivery.

What reason is there to expect chemistry to be able to make progress on the many specific problems contained in these two major general classes of problems? A number of recent advances in science and engineering have provided an extraordinarily powerful set of techniques which are appropriate for the investigation of problems in chemical biology at the molecular and supramolecular level. Examples are given below.

- Recombinant DNA technology provides the opportunity to prepare proteins in substantial quantities with controlled variations in amino acid sequence.
- Monoclonal antibodies produced in large quantities using hybridoma techniques provide specific receptors against virtually any high or medium molecular weight substance.
- X-ray single crystal analysis can now solve very complex structural problems.
- Electron microscopy permits direct visualization of many structures at levels between 2 and 1000 A and is superbly suited to solving problems in membrane chemistry and in the structures of large molecules.
- Ultrasensitive analytical techniques (fast-atom-bombardment mass spectrometry, two-dimensional NMR spectroscopy, antibody-based competitive binding assays, electron capture gas-liquid chromatography, high performance liquid chromatography) can elucidate structures of small quantities of very complex substances.
- Sophisticated synthetic technology, based both on conventional chemical techniques and on a range of biological techniques, including enzymatic transformation and fermentation, make it possible to synthesize complex naturally occurring substances and analogs in useful quantities.
- Computers permit analysis of the details of receptor-substrate binding, and modeling of the complex set of coupled reactions which take place in living cells.

In summary, molecular biology has become sufficiently detailed that many of its most important current questions must be phrased and answered using chemical methods. Hence, chemical techniques are essential to understanding biological phenomena and for converting this understanding into a form useful in medicine and agriculture. There is a pressing practical reason for accelerating this application of chemistry to biological problems, namely, the economic and technological relevance of chemical biology to biotechnology and applied biology, and the intense international competition in this area. Failure to encourage the chemical aspect of biology will cost us leadership in this area and we will forfeit ownership of the practical results to other countries.

4. *Collaboration.*

An important aspect of the proposed program would be its emphasis on interdisciplinary research. Where appropriate, the program would actively encourage development of collaborative projects which integrate programs of chemical and biological research.

In addition, biotechnology is an area of extremely active industrial research. This derives from the special potential for health care (with efforts by essentially all pharmaceutical companies and by a host of small start-up biotechnology companies) and improvements in agriculture (Monsanto, du Pont, and many others). The major immediate markets being addressed by these companies are high-value proteins (interferon, plasma thrombinogen activator, insulin, serum Factor VIII, animal and human growth hormones, vaccines) and monoclonal antibody products (clinical diagnostics). Longer-term industrial projects are aimed at a range of more complex health-care products (treatments for cancer, diabetes, high blood pressure; methods for organ transplantation; immuno-regulators), improved plant strains, enzymatic catalysts for industrial and food use, biomass utilization for energy, and human genetic engineering.

The ultimate success of this industrial endeavor will depend heavily upon the supply of young scientists with the needed cross-disciplinary expertise. This program will provide such experience through research more basic than but clearly complementary to the major threads of current industrial research. For example, the development of molecular understandings of the reactions in nerve transmission and of oxidative phosphorylation are central to advances in the relevant biology but still far from application. Ultimately, these areas will be central to the development of neuroactive drugs and more efficient plants. Similarly, the development of new synthetic methods for biologically relevant molecules remains primarily an academic activity, but it is critical to the continued leadership of our pharmaceutical industry.

In fact, the scope for new research in chemical biology is so great that full and cooperative effort from both university and industry is both warranted and required. Industry is now actively building on the basic science of recent years (as well as conducting some of its own very excellent basic work); the university must provide the basic science for the next cycle of industrial discovery and development.

III. CONCLUSIONS

Chemistry extends all the way from physics, with its strong ethos of deductive logic, to the largely phenomenological biological sciences. We are at a gratifying period when chemistry is "hot at both ends." With adequate access to the powerful techniques now developing, chemistry can expect to explore and understand the principles that govern chemical change at a much more fundamental level than has been possible in the past. At this end of the spectrum, chemistry can begin to escape the bounds of empiricism imposed thus far by the intractability of many-body theory and the limits of our measurements. Yet, despite these barriers, chemists have been so enormously successful with their inductive approach that molecules of exquisite complexity can be recognized and synthesized in every intricate detail. This capability is most timely, because the phenomenological advances of the biosciences now demand explication at the molecular level. Thus, the time is ripe for quite spectacular advances in chemistry.

Hence, in the next ten to twenty years, there will be dramatic changes in our basic understandings of chemical change and our ability to marshal those understandings to deliberate purpose. The program presented here is intended to define a leadership role for the U.S., as these advances are won. The rewards accompanying such leadership are commensurate with the prominent role of chemistry in addressing societal needs, in ameliorating the problems of our technological age, and in sustaining our economic well-being. The costs of falling behind are simply intolerable.

To exploit the intellectual opportunities in chemistry and to maximize the social gain, a substantial incremental increase in federal support for basic research in the chemical sciences is needed. This need for enhanced federal investment in chemical research is rooted in a pattern of funding historically appropriate to a test-tube and bunsen burner era, an era long since eclipsed. The sophistication of a modern chemistry laboratory requires a much more vigorous and sustained financial commitment, both in capital investment and in supporting services. The cost is miniscule compared to the stakes involved; we must nurture a $175 billion industry that maintains a $12 billion positive balance of trade and employs over a million people. We must continue to provide it with a full reservoir of fundamental knowledge.

Appendix I

Fundamental Questions Before Us

A. UNDERSTANDING CHEMICAL REACTIVITY

Molecular Dynamics.

a. How is energy redistributed as a function of time if it is deposited in a molecule in a specific excitation?
b. How does the energy in a molecule (amount, states, intramolecular distribution) determine its reactivity?
c. Can the chemistry of a molecule be affected by an intense (but not absorbed) photon field?
d. Can we understand and theoretically predict the detailed energy distribution among the products in bimolecular reactions?
e. Can we observe the transition state in real time?
f. Is there a Franck-Condon aspect to H or H+ transfer?
g. How can a non-thermalized distribution of excited states be maintained and can novel chemistry be achieved thereby?
h. Can we fully characterize and understand intramolecular and intermolecular energy transfer processes?
i. Can we store solar energy using artificial photosynthesis?
j. What new reaction mechanisms operate at high temperatures because of unusual gaseous molecules and unusually rapid mobility?
k. Can we understand and predict the unusual chemical species and solid phases formed at high temperatures?

Chemical Pathways.

a. What are the systematics of the metal-ligand bond energies in organometallic compounds?
b. What are the mechanisms and rates of multielectron transfer reactions and how can they be controlled?
c. Can interfacial energy be used to synthesize new solids with unique structural features (e.g., channels, cavities, holes)?
d. How can metastable non-stoichiometric solids be prepared?
e. How can the dimensionality of a solid be controlled (linear chains, two dimensional nets, sandwich solids, etc.)?
f. What principles govern conformer-selective chemistry?
g. Can synthetic polymers that are dynamic entities be designed?
h. Can novel inorganic solids be made with molecular beam epitaxy?
i. Do topological principles underlie some aspects of chemical structures and reactions?
j. What molecular properties exhibit bistability suitable for digital information storage in a molecular electronics system?
k. Can organic ferromagnetic compounds be synthesized?
l. What are the structural requirements for nonlinear optical properties in organic molecules?

B. CHEMICAL CATALYSIS

Heterogeneous Catalysis.

a. For chemisorbed intermediate species, what are the residence times? Can they be changed to raise catalyst efficiency?
b. For chemisorbed intermediate species, what molecular structures and oxidation states give optimum selectivity?
c. Why are transition metals the best catalysts?
d. What general mechanisms determine catalyst deactivation? Specifically, how do nickel and vanadium compounds deactivate catalysts?
e. What is the combustion process of carbon on catalytic surfaces?

f. What determines surface migration of chemisorbed species on a surface? What role does this migration play in catalysis?
g. Can we identify and characterize active sites in heterogeneous catalysts?

Homogeneous Catalysis.

a. How is the rupture of H-H, C-H, and C-C bonds altered as the size of a metal cluster catalyst increases?
b. What relationships connect homogeneous, metal cluster, and metal surface catalysis?
c. Can we design homogeneous catalysts that would selectively convert coal to specific desired products?
d. How can chiral metal complexes be designed to induce chirality in organic reaction products?
e. Can we achieve selective homogeneous oxidation catalysis for terminal positions in alkanes?
f. Can catalytic Diels-Alder chemistry be developed?
g. Can oxidative fixation of nitrogen (e.g., to nitric acid) be catalyzed?
h. Can the addition of water or ammonia to olefins (to give alcohols and amines) be catalyzed?

Photocatalysis and Electrocatalysis.

a. How does catalysis at the solid-gas interface connect to electrocatalysis at the solid-liquid interfaces?
b. Can we define, control, and understand the electronic structure of surfaces?
c. What principles are involved in the design and synthesis of new electrocatalysts?
d. What principles are involved in the design and synthesis of new semiconductors materials for photoassisted reactions?
e. Can we develop analytical methods to study chemistry at the solid/liquid interface to match the power of those recently developed for use at the solid/vacuum interface?
f. How can light-absorbing dyes be chemically bound to semiconductor surfaces?

Artificial-Enzyme Catalysis.

a. What obstacles impede synthesis of artificial-enzyme catalysts tailored for a desired reaction?
b. Can we design catalysts to fix the structure (head-to-tail, head-to-head, block, etc.) and stereochemistry of polymers?
c. Can we enhance selectivity through surface shape?
d. Can artificial enzymes be designed to operate in organic solvents?
e. Can artificial enzymes be incorporated into artificial membranes to achieve energy-driven transport (as occurs in living cells)?
f. By adding binding sites for regulator molecules, can we design artificial enzymes to act as detectors? Can these be developed with feedback loops to adapt and control their environment?

C. CHEMISTRY OF LIFE PROCESSES

Enzymology.

a. What is the relation between receptor protein/enzyme binding/active site structure and specificity/catalytic activity?
b. What strategies are required to design substrates which compete with endogeneous substrate for specific receptor/active sites?
c. Can enzymes be used as practical catalysts in organic synthesis?
d. What are the mechanisms of enzymatic reactions (e.g., in protein and nucleic acid synthesis, oxidative phosphorylation, photosynthesis, active transport, intermediary metabolism)?
e. Can combined chemical/biological synthetic techniques be developed to improve markedly the synthesis of nucleic acid sequences?
f. What is the atomic-level basis for enzyme tertiary structure and stability?

Immunology.

a. What is the relation between immunoglobin structure and binding specificity?

b. What are the best techniques for producing and purifying monoclonal antibodies?
c. What antigens and adjuvants are most effective in initiating an immune response?
d. How large a fragment of a protein antigen is required to evoke an effective immune response?
e. Can monoclonal antibodies serve as a basis for chemical purification procedures? As starting materials for semisynthetic enzyme production? For new types of chemical analytical schemes?

Endocrinology.

a. What are the structures of important endogenous hormones?
b. What is the best strategy for inferring the binding specificity of hormone receptors?
c. How should one design synthetic hormone analogs that have desired activity, specificity, and *in vivo* lifetime?
d. What synthetic strategies are required to synthesize synthetic hormones and analogs (especially polypeptide hormones)?

Neurochemistry.

a. What are the structures of important neurotransmitters?
b. What related structures will have agonist/antagonist activity?

Membrane Chemistry and Vectorial Chemistry.

a. Is the chemiosmotic hypothesis correct? What is the molecular basis for interconversion of energy stored in concentration gradients and energy stored in reactive chemical bonds?
b. How is charge separation across membranes accomplished in photosynthesis, respiration, and related areas of bioenergetics?
c. What is the molecular basis for active transport processes? Can understanding of these processes be used to suggest new types of abiological membranes having useful selectivity?
d. What is the chemistry of perception (vision, smell, memory)?
e. What is the relation between chemical, physical, and biological properties of lipid membranes? Can synthetic membrane systems (liposomes, others) be developed for drug delivery?

Chemical Studies Relevant to Biological Systems.

a. Can we understand the hydrophobic effect in aqueous solutions?
b. What are the best strategies for the synthesis of water-soluble molecules having biological activity (oligosaccharides, lipids, polypeptides, neuroactive substances, nucleic acids, metabolites)?
c. What are the structural and thermodynamic bases for molecular recognition (host/guest chemistry)?
d. What are the properties of thin hydrocarbon membranes (selectivity, permeability, concentration polarization, interfacial structure, thermodynamics of membrane diffusion)?
e. How are cooperative chemical phenomena related to biological systems (liquid crystals, organized lipid assemblies, oscillating reactions, nucleation and phase transitions)?
f. Can one construct instructive chemical models for important biological phenomena (active transport, photosynthesis, oxidation phosphorylation, proton translocation, muscular contraction, information storage, nerve conduction, self-replicating systems)?

RESEARCH BRIEFINGS 1984

for the Office of Science and Technology Policy,
the National Science Foundation,
and Selected Federal Departments and Agencies

Committee on Science, Engineering,
and Public Policy

National Academy of Sciences

National Academy of Engineering

Institute of Medicine

NATIONAL ACADEMY PRESS
Washington, D.C. 1984

National Academy Press 2101 Constitution Avenue, NW Washington, DC 20418

NOTICE: The National Academy of Sciences was established in 1863 by Act of Congress as a private, nonprofit, self-governing membership corporation for the furtherance of science and technology for the general welfare. The terms of its charter require the National Academy of Sciences to advise the federal government upon request within its fields of competence. Under this corporate charter, the National Academy of Engineering and the Institute of Medicine were established in 1964 and 1970, respectively.

The Committee on Science, Engineering, and Public Policy is a joint committee of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. It includes members of the councils of all three bodies.

This work was supported by the National Science Foundation under Contract PSPF91304.

Library of Congress Catalog Card Number 83-63234

International Standard Book Number 0-309-03495-7

Copyright © 1984 by the National Academy of Sciences

No part of this book may be reproduced by any mechanical, photographic, or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted, or otherwise copied for public or private use, without written permission from the publisher, except for the purposes of official use by the United States Government.

Printed in the United States of America

Committee on Science, Engineering, and Public Policy

LEON T. SILVER, Professor of Geology, Division of Geological and Planetary Sciences, California Institute of Technology (*Chairman*)
LINDA H. AIKEN, Vice President for Research, The Robert Wood Johnson Foundation
RICHARD C. ATKINSON, Chancellor, University of California, San Diego
JACOB BIGELEISEN, Leading Professor, Department of Chemistry, State University of New York, Stony Brook
FLOYD E. BLOOM, Director and Member, Division of Pre-Clinical Neuroscience and Endocrinology, Scripps Clinic and Research Foundation
W. DALE COMPTON, Vice President-Research, Ford Motor Company
EMILIO Q. DADDARIO, Wilkes, Artis, Hedrick and Lane, Attorneys at Law
GERALD P. DINNEEN, Vice President, Science and Technology, Honeywell, Inc.
ALEXANDER FLAX, President Emeritus, Institute for Defense Analyses
GARDNER LINDZEY, President and Director, Center for Advanced Study in the Behavioral Sciences
†GEORGE M. LOW, President, Rensselaer Polytechnic Institute
EDWARD A. MASON, Vice President-Research, Standard Oil Company (Indiana)
*JOHN L. McLUCAS, Executive Vice President and Chief Strategic Officer, Communications Satellite Corporation
*ELIZABETH C. MILLER, Van Rensselaer Potter Professor of Oncology, McArdle Laboratory for Cancer Research, University of Wisconsin
DANIEL NATHANS, Professor, Department of Molecular Biology and Genetics, The Johns Hopkins University School of Medicine
GILBERT S. OMENN, Dean, School of Public Health and Community Medicine, University of Washington, Seattle
HERBERT A. SIMON, Professor of Computer Science and Psychology, Department of Psychology, Carnegie-Mellon University

† Past Chairman, deceased July 17, 1984
* Term expired June 30, 1984

*I. M. SINGER, Professor, Mathematics Department, University of California, Berkeley

F KARL WILLENBROCK, Cecil H. Green Professor of Engineering, School of Engineering and Applied Science, Southern Methodist University

Ex Officio

FRANK PRESS, President, National Academy of Sciences

ROBERT M. WHITE, President, National Academy of Engineering

FREDERICK C. ROBBINS, President, Institute of Medicine

COSEPUP Staff

ALLAN R. HOFFMAN, *Executive Director*

BARBARA A. CANDLAND, *Administrative Assistant*

---

*Term expired June 30, 1984

Preface

This volume contains nine research briefing reports prepared in 1984 under the supervision of the Committee on Science, Engineering, and Public Policy (COSEPUP), a joint committee of the National Academies of Sciences and Engineering, and the Institute of Medicine. Prepared at the request of the Office of Science and Technology Policy (OSTP) and the National Science Foundation (NSF), the briefings were delivered in both written and oral form to these organizations between September 5 and October 1, 1984. The briefings were also presented to senior officials of other interested federal departments and agencies.

Potential briefing topics were identified by OSTP, the Academies, the Institute of Medicine, and units of the National Research Council. Final selections were made by OSTP. It must be emphasized that presentation of these nine briefings in 1984 by COSEPUP does not imply that other topics not presented are considered to be of lesser importance to the progress of U.S. science and technology. No priority setting among fields of science and technology by COSEPUP is implied. The list reflects the fact that only a limited number of timely topics can be covered in a given round.

The nine briefing topics for 1984 are:

1. *Computer Architecture*: Research directed toward increasing the speed and capabilities of computers through innovative hardware design and software optimization, including parallel processing.

2. *Information Technology in Precollege Education*: Research directed toward identification of special opportunities for use of new information technology in precollege education offered by recent work, notably in the cognitive sciences, artificial intelligence, hardware, and analyses of the strengths and limits of currently available and anticipated devices.

3. *Chemical and Process Engineering for Biotechnology*: Long-range fundamental research directed toward developing new reactor concepts (immobilized-enzyme reactors and continuous large-scale cell culture systems) and developing new unit operations for manipulating and purifying fragile macromolecules in highly complex product mixtures.

4. *High-Peformance Polymer Composites*: Research directed toward synthesis of high-strength, high-modulus polymers and polymer composites for structural applications, including examination of relationships between microstructure and performance within single-phase materials, and relationships between surface characteristics and interfacial properties in multiphase composites.

5. *Biology of Oncogenes*: Research directed toward enhanced understanding of the genetic characteristics of cellular and viral oncogenes, the process of activation of oncogenes and its relationship to chemical/environmental carcinogenesis, and the nature of oncogene products and how they subvert normal orderly cellular growth and development.

6. *Interactions Between Blood and Blood Vessels (Including the Biology of Atherosclerosis)*: Research directed toward enhanced under-

standing of the interplay between blood constituents and the blood vessel wall, ranging from molecular and cellular interactions to the role of injury and subsequent disease.

7. *Biology of Parasitism*: Research directed toward enhanced understanding of how parasites establish and maintain relationships with their hosts and the consequent disease processes, including examination of:

- cell membrane phenomena;
- intermediary metabolism;
- life cycle and development changes in vectors and in hosts;
- immune system and other protective responses and their underlying genetic and regulatory mechanisms; and
- related opportunities to develop vaccines, pharmaceuticals, and biological controls.

8. *Solar-Terrestrial Plasma Physics*: Research directed toward understanding the interconnected physical processes responsible for the injection, transport, energization, and loss of plasma from regions on the sun, in the solar winds, and in the earth's magnetosphere.

9. *Selected Opportunities in Physics*: Research directed toward enhanced understanding of the following promising areas of physics: physics at the laser-atomic frontier; relativistic plasma waves; physical properties of deliberately-structed materials; biomolecular dynamics and intercelluar cooperativity; cosmology; and nuclear matter under extreme conditions.

This set brings to 21 the number of research briefings developed for OSTP and NSF by COSEPUP in three separate rounds. Those topics for which briefings were developed in 1982 and 1983 are listed below:

1982 (Published as *Research Briefings for the Office of Science and Technology Policy , the National Science Foundation, and Selected Federal Departments and Agencies*, National Academy Press, 1983).

1. Mathematics
2. Atmospheric Sciences
3. Astronomy and Astrophysics
4. Agricultural Research
5. Neuroscience
6. Human Health Effects of Hazardous Chemical Exposures
7. Materials Science

1983 (Published as *Research Briefings 1983 for the Office of Science and Technology Policy , the National Science Foundation, and Selected Federal Departments and Agencies*, National Academy Press, 1983).

1. Selected Opportunities in Chemistry
2. Cognitive Science and Artificial Intelligence
3. Immunology
4. Solid Earth Sciences
5. Computers in Design and Manufacturing

Each briefing is developed by a panel of approximately 12 members, charged with critically assessing its field, and identifying those research areas within the field likely to return the highest scientific dividends as a result of incremental federal investment in the upcoming fiscal year. Panels generally meet once, for several days, to carry out their charge, with rapporteurs present to summarize the discussions and prepare initial drafts of briefing papers. After further review by the panels, the briefing papers and oral presentations are reviewed by COSEPUP, prior to their delivery to federal officials in three separate briefings. Funds to support this round, as well as both earlier rounds of research briefings, have been provided by the National Science Foundation.

Publication of this volume comes shortly after the untimely death of COSEPUP's past Chairman, George M. Low, under whose capable leadership the research briefing activity was initiated and developed. George's guidance and wisdom will be sorely missed.

Leon T. Silver, *Chairman*
Committee on Science, Engineering, and Public Policy

Contents

Report of the Research Briefing Panel on Computer Architecture 1

Report of the Research Briefing Panel on Information Technology in Precollege Education 17

Report of the Research Briefing Panel on Chemical and Process Engineering for Biotechnology 31

Report of the Research Briefing Panel on High-Performance Polymer Composites 45

Report of the Research Briefing Panel on the Biology of Oncogenes 57

Report of the Research Briefing Panel on Interactions Between Blood and Blood Vessels (Including the Biology of Atherosclerosis) 69

Report of the Research Briefing Panel on the Biology of Parasitism 81

Report of the Research Briefing Panel on Solar-Terrestrial Plasma Physics 91

Report of the Research Briefing Panel on Selected Opportunities in Physics 105

Report of the Research Briefing Panel on Computer Architecture

Research Briefing Panel on Computer Architecture

Jacob T. Schwartz (*Chairman*), New York University, New York, N.Y.
Richards Adrion, National Science Foundation, Washington, D.C.
Donald Bonstrom, ETA Systems, St. Paul, Minn.
Charles Bostick, National Security Agency, Fort Meade, Md.
John Cavalini, Department of Energy, Washington, D.C.
Robert J. Douglass, Los Alamos National Laboratory, Los Alamos, N. Mex.
W. Daniel Hillis, Thinking Machines Corporation, Waltham, Mass.
Thomas A. Keenan, National Science Foundation, Washington, D.C.
David Kuck, Kuck and Associates, Champaign, Ill.
H. T. Kung, Carnegie-Mellon University, Pittsburgh, Pa.
Thomas Nash, Fermi National Accelerator Laboratory, Batavia, Ill.
James Pomerene, IBM Corporation, Yorktown Heights, N.Y.
John Riganati, National Bureau of Standards, Gaithersburg, Md.
Arnold Reisman, Microelectronics Center of North Carolina, Research Triangle Park, N.C.
Paul Schneck, Office of Naval Research, Arlington, Va.
Stephen Squires, DARPA/IPTO, Arlington, Va.
Kenneth Wallgren, NASA, Washington, D.C.

Staff

Mark Weiser, *Rapporteur*, University of Maryland, College Park, Md.
Gary E. Clark, *Staff Director*, Computer Science and Technology Board
Rosena A. Ricks, *Administrative Secretary*, Computer Science and Technology Board
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

Report of the Research Briefing Panel on Computer Architecture

I. INTRODUCTION

The market for supercomputers is only a small part of the total computer market. However, it represents a critical cutting edge for the advance of computer technology and provides machines that are critical to technological leadership in a whole range of areas varying from aircraft design to basic chemistry. For this reason it is vital that the programs that achieved U.S. supercomputer leadership be continued and that new programs to maintain that leadership be initiated.

Rapidly developing VLSI technology has created the basis for major advances in superspeed computer performance. Maximum attainable computation rates, which are approaching 1,000 MIPS (millions of instructions per second) for the largest vector supercomputers, will rise over the next decade through 20,000 MIPS, and possibly to levels of 100,000 MIPS and more. Attaining such extreme speeds will depend on the effective use of parallelism, i.e., on the development of computers that can execute many hundreds, thousands, or tens of thousands of instructions simultaneously and of software that can effectively orchestrate many simultaneous streams of computation. The ability to explore this essential new opportunity has been opened up by the amazing reductions in circuit costs attained during the past decade.

This document will review the current supercomputer situation and will predict developments over the next 5 to 10 years. The following major areas will be considered:

1. the *chip and mass storage technologies* on which U.S. primacy in the computer area, and its ability to advance, ultimately rest;
2. the *computer architectures*, both novel and relatively conventional, likely to be important for large-scale computing over the next decade;
3. the potentially significant role of *special-purpose computers*, which may be of particular importance for certain types of research (e.g., computer vision, speech and signal processing, scientific data reconstruction and reduction), where dedicated equipment and very high performance, but relatively inexpensive equipment, are all essential;
4. the extent to which *software technology* can and must support the move to large-scale parallel computing;
5. the developments in *communication tech-*

nology that will define the ways in which supercomputers can be made available to the U.S. research community;

6. the *efforts abroad* that might undercut the dominant U.S. role in high-performance computing;
7. the relationship of supercomputing to the broader field of *artificial intelligence*; and
8. the bottom-line technical policy question of how to get the U.S. research and industrial communities to *move forward in rapid and effective cooperation*.

II. CHIP TECHNOLOGY

The astonishing exponential growth of microchip performance is both the main driving force for the computer industry and the technical basis for expected parallel supercomputer developments. Though many circuit families of differing speeds and cooling requirements play a role in this growth, the following remarks will concentrate on just one family, the so-called field-effect transistors (FETs), which generate 40 times less heat than the faster bipolar chips, a very significant advantage as the number of circuits per chip increases steadily. State-of-the-art FET chips as of 1984 are represented by the 256K memory chip (512,000 devices), just now coming into mass production, and the roughly 100,000-device Motorola 68000 microprocessor. The 5 to 1 circuit density advantage of memory chips results from their very regular internal geometric structure and highly optimized design. Since 1960 these circuit densities have doubled roughly every 18 months, and this growth is expected to continue, with memory and microprocessor densities remaining in a similar proportion. By 1988 this should make 2-million-bit/chip memories available and could make million-transistor single-chip processors possible. As circuit sizes shrink, speeds will rise, so by 1988 the cycle-time of single-chip FET processors should have fallen from the 80 nanosec now typical to 30 nanosec for chips operated at room temperature and 15 nanosec for chips operated cryogenically. A typical application of this technology might be a single-chip, 64-bit scientific processor with full floating-point capability, on-chip memory management, internal overlap of instruction execution, 32 Kbytes of on-chip cache memory, and the capability of performing 20-40 MIPS. Processors tailored for somewhat faster execution of higher-level languages such as LISP and special-purpose computing devices like those described in Section V below (e.g., specialized vision-oriented or switching chips) all can be supported by this same technology.

The use of new materials such as gallium arsenide and other techniques for attaining higher electron mobility should allow perhaps 10 times higher performance, but they require significantly more difficult process and materials technology. The Japanese HEMT and the U.S. VHSIC are examples of such projects. Josephson circuits, which operate at cryogenic temperatures, and 3-D chips, which try to build up circuits in the third dimension rather than using the essentially planar structures now common, are not likely to be useful technologies in the near future.

Wafer-scale integration, including full and partial wafer techniques, have been considered for reducing the costs associated with the currently used wafer-dicing and reassembly technology. This approach faces severe chip yield problems, and one must reckon with the unfavorable signal-propagation characteristics of long on-chip conductors. For this reason, a packaging technique like that used in the IBM Thermal Conduction Modules, i.e., bonding individual chips to a multilayer ceramic substrate, may dominate for high-performance devices. Assuming that 200-400 chips are bonded to a single 6" × 6" ceramic board of this type, two such boards might by 1990 produce a 64-processor, 1,200 MIPS peak-rate parallel computer at a manufacturing cost between $50 thousand and $100 thousand.

Attaining the chip performances described in the preceding paragraphs will require attention to a challenging mass of supporting technologies, including special materials development, x-ray, E-beam, and ion lithography, as well as radiation damage removal methods. This work is moving rapidly in the United States and at least as rapidly in Japan, where experimental and fabrication facilities sometimes bettering the best U.S. facilities have been put in place. Japanese work in this critical area is spurred by a particularly clear national determination to succeed, resulting in a willingness to try approaches that U.S. workers tend to avoid because of initially unclear market perspectives. In this determination, combined with the dedicated large teams and engineering excellence brought to bear, lies the technological roots of the Japanese computer industry's rapid advance.

III. DISK TECHNOLOGIES

For supercomputers to be used effectively, they must be supplied as balanced systems, with input-output (I/O) bandwidth, external storage, displays, and other capabilities all appropriately matched to the expected two- or three-orders of magnitude increase in computation rates. Concern has been expressed about the ability of storage disks, a key element of this "peripheral" technology, to keep up with such rapid speed increases. The workhorse disk for current supercomputer systems is still the CDC Model 819 disk, a rotating storage device dating from 1970, with a capacity of 600 million bytes and a 5-Mbyte-per-second transfer rate. Two newer disks just now becoming available provide roughly double this capacity and storage rate, an exceedingly modest increase compared to far more rapid computation-rate speedups. Other more ambitious magnetic disk alternatives include the possible use of synchronized spindles or multiple heads in parallel. Although no such higher performance products have been announced (probably because of less-than-satisfactory past market experience with very-high-performance disks), products of this type may appear as derivatives of disk and head technology developed for larger, more lucrative markets.

Against this background, the development of a new optical disk storing 10^{11} bytes, now in progress under NASA and USAF (RADC) sponsorship, appears extremely significant. This disk, a first version of which should be available in 1987, promises to support a 10^9 byte/sec transfer, much better matched to anticipated supercomputer requirements than anything currently available. Access time, i.e., the time needed to start the inflow of a randomly selected data record, is anticipated to be 50-100 milliseconds. The development of this potentially important device ought to be pushed rapidly to prevent an initially limited market from delaying its appearance.

IV. CURRENT AND NEW COMPUTER ARCHITECTURES

During the next 5 years, new types of machines based on large-scale parallelism, having the ability to perform thousands or even tens of thousands of arithmetic operations simultaneously, will begin to appear alongside high-speed sequential and vector computers that dominate today's high-end computer market. At a rate determined by their performance, their cost, and their suitability for various applications, these new machines may progressively supplant present designs in some market sectors. Though our panel did not make any precise predictions concerning the rate at which this development will unfold over the next decade, we can describe the main architectural considerations relevant to this process.

High-speed single-thread machines keep programming relatively straightforward by behaving externally as if they execute instructions serially, even though internally a small number of operations (5-20) are actually being handled simultaneously. Machines of this design are the staple of present medium-high performance (4-20 MIPS) computing. It is technically possible to raise the performance of these machines to the 80-150 MIPS range with current technologies and to 250 MIPS or somewhat more using more advanced technology such as that represented by GaAs. Performance improvements beyond this will depend on the parallel use of multiple high-speed serial processors. Very-high-speed serial machines and multiprocessor configurations involving small numbers of such computers are bound to retain major, perhaps dominant, economic importance for the next 5-10 years, especially for commercial applications. In the supercomputer environment, fast (4-10 MIPS) serial computers will be the single-user "workstations" that supply researchers with their software development tools and much of their graphic and data display environments.

Vector supercomputers, like the present U.S. Cyber 205, CRAY XMP, and the Japanese Hitachi S-810/20 and Fujitsu VP-200, can be regarded as simple parallel machines specialized for the handling of computations characterized by particularly regular patterns of data motion. Such regularities typify certain important numerical scientific codes, and hence fast vector machines are generally designed as "scientific number crunchers" equipped with the highest-speed floating-point operation units. The current peak performance of these machines is in the hundreds of MIPS. Single-processor machines of this type could achieve operation rates of 250-500 MIPS with the best present technology and perhaps twice this using GaAs technology. Multiprocessor configurations of vector processors, a type of design under active consideration by several supercomputer manufacturers, will probably attain peak performances of roughly 10,000 MIPS[*] using 16-32 processors within 5 years, rising to 25,000 MIPS within the coming decade. At such performance levels, they can be expected to retain a major role, and perhaps a dominant role, in bread-and-butter scientific number crunching through much of the decade. Note, however, that computers of this design can exploit parallelism only in innermost program loops; and for programs that branch or address memory in irregular patterns, the actual efficiency typically decreases to 15-20 percent of rated peak performance, e.g., for such problems a 4,000 MIPS actual computation rate is all that can be expected from a machine nominally rated at 25,000 MIPS.

New parallel supercomputer designs, many of which are still highly experimental, are at once the focus of much current research and the one area (of design, as distinct from technology) in which sudden surprising advances might occur. Many of these designs promise large increases in computational power without significant redesign of their components, i.e. they are "scalable." With such architectures it may be possible to increase machine speeds by factors of 10 or even 100 by increasing the amount of parallelism used, without major redesign but at a proportional increase in cost. Once suitable software has been developed for machines of this type, a major research undertaking, the homogeneity of their structure should make it possible to move applications to much larger-scale parallel machines with minimal software change. Moreover, scalable parallel machines have major potential advantages in a VLSI environment because they are built from large numbers of identical parts that can be mass produced efficiently and because the design cost can be amortized over multiple configurations differing only in size. For all these reasons, we review this class of computers at length and attempt to clarify the technical options available for parallel supercomputer development by listing various principal design alternatives characterizing the design of such machines.

Most of the more than 70 parallel machines that have been proposed by university and

[*] The performance of high-speed scientific computers is often rated not in terms of the total rate at which instructions of all types are performed but only counting that part of the total activity devoted to the execution of "floating-point" arithmetic operations: MFLOPS (millions of floating-point instructions per second) rather than MIPS (millions of instructions per second). The ratio between these two measures is typically taken to be 1/3, but since this ratio is highly application dependent, we have chosen to sidestep this distinction and to state all crude performance figures in MIPS.

industrial groups here and abroad occupy easily characterized positions in the "design space" defined by the following alternatives:

1. *SIMD versus MIMD*. An SIMD (single-instruction multiple-data) machine consists of multiple processors (or of a few very fast pipelined processing elements), all of which are fed synchronously by a single instruction stream but which operate on independent streams of data. Vector machines like the CRAY I or Hitachi 810, and synchronous array machines like the ILLIAC IV and NASA/Goodyear MPP, typify this class. An MIMD (multiple-instruction multiple-data) computer consists of multiple processors (or fast time-sliced processing elements) driven by independent instruction streams and capable of branching independently but also able to pass data and synchronization signals between processors, perhaps via a shared memory. The Denelcor HEP, Cal Tech Homogeneous Hypercube, and numerous other machines belong to this class. SIMD machines are generally more efficient for applications characterized by very regular patterns of processing, in part because their "lockstep" mode of operation shares a single unit of control hardware among many data-processing elements and eliminates much of the message-passing overhead needed when all the processors of an MIMD machine must proceed through a computation in close synchrony. On the other hand, it is hard to make SIMD computers deal effectively with highly "branched" code, since the process of branching in such computers is normally imitated by temporarily disabling some of the processors in the parallel array, and thus several successive branches tend to reduce effective computational power drastically. MIMD processing arrays deal more robustly with complex branched code, since they are capable of executing not only innermost loops, but even complex branch-filled loops, in parallel.
2. *Visible versus "under the covers" physical communication pattern*. In all parallel machines some portion of the computation involves communication among the individual processing elements. From the physical point of view such communication is allowed in only a few specific patterns defined by the hardware. For example, the processors may be arranged in an obvious two-dimensional grid with each processor connected to its north, south, east, and west neighbor. A single operation on a machine wired in this pattern can only send a number from each processor to one of these neighbors. Proposed connection patterns for parallel machines include rings, n-dimensional cubes, binary trees, and various less obvious but much more powerful communication networks developed by communications researchers.

   Once having chosen a communication net appropriate for the computational structure being developed, a parallel machine designer can choose either to make the structure of this net visible to the machine's user or to hide these details behind a set of programmed conventions that makes it appear that any processor can communicate directly with any other.

   Both alternatives have arguable advantages. On the one hand, if a machine's actual communication pattern is well matched to an application and is made available to the programmer, it can be exploited to gain speed. Examples are the use of two-dimensional grid patterns for image processing and the use of various powerful communication patterns for optimal implementation of the Fast Fourier Transform. On the other hand, machines that relieve their programmer from direct concern for the pattern in which messages must flow between processors are easier to program for a wide range of problems, particularly problems characterized by highly variable patterns of communication. Another potential advantage of machines in this latter class is that the communication paths that they use, being hidden, can be varied dynamically, e.g., to bypass faulty components.

3. *Coarse-grained versus fine-grained designs.* In any parallel computer with multiple processing elements there is a cost tradeoff between the number and the size of the processors. The approach that stays closest to established single-stream processor technology is to use as few as possible of the largest available processors. The opposite approach is to achieve as much parallelism as possible by using a large number of small elementary processors, which can be single-chip microprocessors chosen to exploit the cost-effectiveness of a mass-produced item or, in the limit, can be a single-bit processor used in very large numbers. The decision to use a relatively small number of very-high-speed processors is justifiable by the argument that even largely parallelizable code will have significant serial sections, so high performance in serial code will be important for sustained overall high performance. Its polar opposite, the use of many single-bit processors, can be justified by a desire to push parallel execution to its limit, and by the surprising level of arithmetic power per square millimeter of silicon that one-bit processors attain. Perhaps the most important issue here is one of programming style. Since ordinary serial machines are coarse-grained, the technology for programming coarse-grained machines is understood best. Thus, it is plausible to expect a FORTRAN compiler to optimize code that keeps, say, 16 processing units busy, but not 16,000. On the other hand, if an algorithm is written with parallel processing in mind from the start, it may divide naturally into numerous processes, e.g., the processing of a 1,000 by 1,000 image in which relatively simple, low-precision algorithms must be applied to individual pixels or small pixel groups may fit most naturally on a million processor machine.
4. *Specialization of architecture to an anticipated application or programming style.* Scientific number crunchers are optimized for arithmetic operations on floating-point numbers, while processors intended for more varied symbolic applications must be optimized to handle irregular memory references and branching patterns, fast subroutine transitions, and fast context switching. They must provide run-time type checking and must support efficient execution monitoring during program debugging. Thus in designing a "scientific" processor one will emphasize certain types of hardware that are of little use in a "symbolic" machine, and vice versa. Various designs for large-scale "dataflow" computers, which are specialized in a different direction, illustrate this same general point. Dataflow programmming is organized around a set of fully parallel concepts. A program is regarded as an assemblage of "operators," all poised to fire as soon as they receive all their logical inputs. As soon as it fires, each such operator transmits its output to further operators, for which it becomes one of several possible inputs.

   The innermost interpretation cycle for such a language consists of various operand-routing, instruction-dispatching, and instruction-execution activities, many of which can be overlapped. As compared to a general-purpose parallel assemblage, a cleverly designed layout of dataflow operation nodes in memory and effective instruction dispatch hardware can probably double or triple the efficiency with which a dataflow language executes, while adding only minimally to the total hardware cost of a parallel system. These remarks typify the speed gains achievable by the architectural specialization of parallel machines.

V. SPECIAL-PURPOSE COMPUTING DEVICES

Extremely high computation rates can be attained efficiently by tailoring electronic hardware to algorithms of special importance (e.g., the Fast Fourier Transform) or to the requirements of a particular computer-intensive problem. Devices of this sort are of particular importance to diverse cutting-edge research efforts and to applications where dedicated equipment is essential but dedicated use of multimillion dollar general-purpose computers such as a large vector supercomputer would be prohibitively expensive.

These special-purpose computing devices point to a class of computing devices highly optimized to particular applications, which might be made possible by hypothetical future advances in compiling technology that allow automatic generation of whatever computing device is best suited for any specified application. Though this is far beyond present capabilities, manually designed special-purpose computing systems are of growing significance to defense uses, such as signal processing and image analysis, to high-energy-physics data acquisition and reduction, and to other areas.

It is steadily becoming easier to realize such systems, either by combining relatively standardized high-performance modules into specially wired configurations or by custom gate-array or VLSI designs. Special systems produced in this way tend to reflect end-user performance requirements closely, and hence they uncover important subproblems of the general supercomputer design problem quickly. Moreover, because of the emphasis on cost-effectiveness typically characterizing their design, they encourage industry to improve the effectiveness of more general computing engines supplied commercially. Also, though machines of this sort sometimes involve approaches that address only one specific problem, others affect larger classes of problems in an extremely cost-effective way, making them desirable attachments for general-purpose supercomputer systems. Encouraging a steady flow of such designs, which sometimes advance the effective frontier of computing in the application areas they address, can strengthen the U.S. competitive position significantly.

The sum of these special design efforts has become an important component of computer research that needs to be explicitly recognized and systematically cultivated. For this, an improved infrastructure for rapid and effective development of special-purpose systems is a prime necessity. In particular, projects motivated by special problems need a basic kit of standard modules out of which optimized architectures can be built. This should include low-cost single-board computers, high-density multiport memories, and high-speed switching modules that allow large communications networks tailored to specific applications to be fabricated rapidly. Some of these items are beginning to appear commercially, but further encouragement to industry and systematic attention to standardization issues are appropriate.

To increase the speed with which prototype special-purpose computational designs can be developed for significant experiments, system-level tools of the same quality as the best current VLSI design tools need to be developed. A focused program to develop such tools and make them available to universities, laboratories, and private-sector organizations appears to be an effective way to strengthen the work of special-system designers. Such tools would also facilitate the transfer of successful prototypes into industrially supported products, which is as important for special high-performance computational designs as for general-purpose supercomputer development.

VI. THE ROLE OF COMMUNICATION TECHNOLOGY

Supercomputers communicate three ways: internally, locally (in the same building), and outside (to machines elsewhere in the world). Existing technology for internal connections uses highly parallel buses and is limited primarily by wire length. Existing nearby machine telecommunications technology runs at a maximum of about 200 Mbit/sec, which is faster than most individual computers can accept information. Current outside connections are limited by available nationwide channels to 3 Mbit/sec, and 50 Kbit/sec is more typical.

In the next 5-10 years it is unlikely that the basic internal communication technology will change much. However, it is possible that within 20 years optical connections or wafer-scale integrated circuit technology will have significantly changed the way computers communicate internally.

Local interconnections will be affected by optical technology much sooner. It is possible that within the next 5-10 years commercial optical local networks running at a gigabit/sec rate will become available. This rate will not be accessible to any individual machine but will permit many supercomputers to share a relatively inexpensive network and exchange data at their internal bus rates.

NBS, DARPA, and cooperating international standards organizations have made large-scale networks possible by specifying protocols and standards that promote interoperability of heterogeneous equipment from different manufacturers. However, nationwide network speeds are currently inadequate for the amount of data consumed by supercomputers. Over the next 5 years the NSF ScienceNet project should be in place with high-speed satellite connections linking the major research universities. This will be a significant help to providing access to supercomputers.

Finally, we note that communication security on these nets may become an important issue, and this area deserves careful and early attention.

VII. ALGORITHMIC AND SOFTWARE ISSUES

Effective parallel processing will require a fundamental understanding of computational models, algorithms, programming languages, optimization techniques, processor-memory-I/O relationships, and interfaces to existing computing systems and end users. Though research on parallel algorithms is expanding, the lack of large-scale parallel machines for experimentation orients current work toward the study of individual algorithms rather than on practical applications, systems questions, or large software structures. Moreover, systematic work on large algorithm libraries, such as would quickly develop around major parallel computing installations, is hardly possible in the present atmosphere of largely theoretical studies. Also, although operating systems for parallel machines raise many subtle problems, it is difficult to address these except by working with real parallel systems.

Parallel programming is much more complicated than sequential programming because it deals with simultaneously occurring events involving potentially complex (time-dependent) interactions, so that elusive and nonrepeatable bugs can occur. For this reason, the usability of large parallel machines will be bound up with the development of programming languages that facilitate the expression of parallelism and the development of sophisticated optimizing compilers that improve the execution efficiency of such languages and find implicit parallelism where it is not explicitly expressed.

Hence, although initial experience (e.g., work on the HEP at the Los Alamos Laboratory) indicates that modest extensions to conventional languages will allow important applications to run with high efficiency on certain types of parallel computers, many new programming language concepts need to be investigated as ways for easing the further exploitation of parallelism. These include languages with vector and set operations (e.g., APL), dataflow languages, applicative languages, languages such as parallel LISP and PROLOG that emphasize recursive branching and backtracking processes, and languages in which communicating sequential processes play a central role. It is important both to encourage theoretical study and development of all these classes of languages and also to ensure that they become available for extensive experimental use as soon as substantial parallel computers suitable for them are constructed.

The present state-of-the-art software to support high-speed parallel computation is also represented by vectorizing compilers and more general parallelizing compilers capable of finding certain opportunities for parallelism automatically by analyzing ostensibly serial code. Both Cray Research and Fujitsu Ltd. have produced such compilers for their vector supercomputers, as have CDC, and U.S. research projects at the University of Illinois and Rice University.

While very useful, these systems are not yet able to automatically translate every standard sequential FORTRAN program into a good parallel program. Indeed, this will remain beyond the capabilities of automatic systems in the foreseeable future because the best parallelization usually requires substantial comprehension of the underlying problem domain. However, these systems do provide invaluable aids to the programmer because they permit him to write parallel programs cleanly in a high-level language without cluttering the language with special-purpose constructs.

Present automatic code vectorization techniques need to be evolved to handle programming for parallel multiprocessor supercomputers and to apply to whatever new parallel languages appear most important. Many of today's vectorization techniques are directly applicable; others will need to be adapted substantially or new methods invented. Program development environments tailored to the special requirements of new parallel systems and languages also need to be developed, especially to aid in the debugging of parallel programs. Graphics systems making it easy to deal with the complex outputs that parallel supercomputers can generate will also be needed.

Complex software developments cannot proceed far in a purely "paper mode." For them to flourish, real and relatively large-scale experiments with functioning hardware are essential. Thus, to accelerate the very challenging software research outlined in the preceding paragraphs, it is essential to deploy or provide remote access to experimental parallel machines as soon as they become available. Unconfining access is essential to the major systems software research and development efforts that will be required to ensure effective use of these new machines.

VIII. EFFORTS ABROAD

U.S. supercomputer developments must be appreciated in the context of efforts abroad, primarily those of the powerful and fast-moving Japanese computer industry, and also relative to an accelerating European effort. Our panel's conclusions concerning the Japanese program are as follows. The largest Japanese vector supercomputers (Hitachi S-810/820 and Fujitsu VP-200) have attained parity with, but not surpassed, corresponding U.S. machines such as the CDC Cyber 205 and the CRAY XMP. Japan's stated national commitment to advance in this area, the great engineering strengths of Japanese industry, and the leading position that Japan has won in certain lines of high-speed semiconductor component production will make for a very closely run race in this area. Nonetheless, the United States can at least expect to maintain with rapidly shifting patterns of computation and communication, with great flexibility and at very high speed.

X. HOW TO MOVE FORWARD

Several major federal initiatives aimed at strengthening U.S. supercomputer-related capabilities have been launched and, if pursued along the strongest technical lines and with due regard for the importance of early and effective industrial involvement, can maintain U.S. computer preeminence. Principal among these are the DARPA Strategic Computing Program, the expanded supercomputer research programs of the Department of Energy, and NSF's Supercomputer Access Program. The steady support that NSF has provided for many, indeed most, of the university groups whose work has defined the principal supercomputer architectural alternatives now available has been of fundamental importance and is planned to continue and even to increase modestly.

Since DARPA's architecture program comprises the largest single federal project addressed to supercomputing, its activities will be broadest and most crucial. Work in many areas, including underlying technologies, improved design automation tools, fabrication means for both VLSI and broad-level special systems construction, conceptual exploration of new special general-purpose parallel computer systems, various small-scale prototyping efforts, and a few critical large-scale prototyping efforts, will all be sponsored.

Our panel is supportive of DARPA's intent to organize the selection of the culminating major efforts of this multiyear program via an orderly series of phases, beginning with conceptual studies aimed at the discovery of significant new computer architectures, progressing to the exploration of those architectures by simulation and experimental software development, then to small-scale hardware prototyping where justified, and finally, in the most promising cases, to the construction of large high-performance prototypes. We note that for these large culminating prototypes, research success, and even more the ability to move successful research results into broad U.S. use before foreign competition follows suit, will clearly be dependent on serious industrial involvement. The proper use of academic resources is in research and small-prototype development; the realization of large systems requires the skills of proven engineering organizations with established track records. For these reasons, we recommend that both DARPA and DOE encourage or require substantial industrial involvement (by either established or start-up companies) in the development of each experimental computer system chosen for large-scale prototyping. We understand that it is DARPA's intent to encourage this type of industrial involvement and we support this decision. Ideally, this should involve a substantial commitment of capital and crucial technical skills by the company (or companies) cooperating in such a development, to ensure that careful and independent private-sector technical judgment concerning project soundness is brought to bear in each such case. Moreover, to strengthen university-industry links and overcome what might otherwise be substantial familiarization delays in moving promising designs into large-scale prototyping, it is recommended that both DARPA and DOE encourage industrial involvement even in the initial smaller-scale stages of parallel computer prototyping. Early industrial involvement in special-purpose computer developments is also desirable.

Though considerably smaller than DARPA's effort, the DOE supercomputer program has maintained a significant focus in areas complementary to, and synergistic with, the DARPA effort. Historically, DOE has been the leading agency for the applications of large scientific computing and has maintained particularly close relationships with the supercomputing community. Since optimization of computer architectures for floating-point computation and for the regular patterns of data manipulation characterizing many numerical applications may demand significantly different designs than are optimal for nonnumerical applications, it is important to maintain a strong program of architectural research and development with a clear scientific-computation focus. DOE's strong tradition and high level of expertise in scientific supercomputing make it appropriate for DOE to play a significant funding role for research on advanced computer architectures aimed at high-performance numerical computing, and acceleration of its program in this area is desirable. DOE is also in the process of significantly expanding its supercomputer access program to include scientists in High Energy and Nuclear Physics, Basic Energy Sciences, and Biological and Environmental Research. This expansion will contribute substantially to solving the national supercomputer access problem. In addition, DOE has acquired several commercial parallel computers and is making them available to many researchers. These DOE programs have played and will continue to play a very important role in maintaining overall supercomputer application capabilities.

As currently defined, NSF efforts will concentrate on making supercomputer capabilities broadly available to the U.S. scientific community, including but not restricted to computer scientists. Though initially this program will concentrate on providing access to existing vector supercomputers, our panel sees the NSF Access Program as having broad continuing implications and recommends that NSF plan to make all major experimental parallel computers nationally accessible as soon as feasible. Efforts of this sort, which serve to build supercomputer-use capabilities, strengthen the U.S. position by exploiting one of our most characteristic advantages, the great breadth and diversity of the U.S. computational community. However, the resources of this community will be heavily taxed by the deep algorithmic and software problems associated with the new classes of machines that will be developed over the next decade. Training of an expanded population of workers in this field must therefore be seen as an essential component of a balanced national effort.

The NSF Supercomputer Access Program is well calculated to serve the requirements of the large U.S. scientific community for whom computers are a tool for other research goals. However, in view of the great importance that NSF basic computer science support has had in the development of new architectural alternatives and the related computer science until now and in view of the excellent technical judgment with regard to hardware and software that NSF can bring to bear, the NSF programs should also be structured to support major software and applications development for the machines to which access is provided. Moreover, to avoid dissipating the attention of research groups among overly many new machines, all the agencies sponsoring supercomputer research and development should plan to support groups wishing to develop algorithms and software for experimental hardware systems produced commercially or by other universities. Such participating research groups should be guaranteed adequately "privileged" remote access, e.g., the right to take full control of a system for scheduled blocks of time to allow exploration of operating system changes in the light of new research ideas. In some cases, it will also be appropriate to disseminate smaller-scale versions of major experimental systems to several locations.

The need for quantitative measurements of progress and quantitative characterization of the advantages of alternative architectural approaches underlies all the areas of research we have discussed. It is therefore appropriate to encourage joint industry-university-government development of tools and techniques to measure progress in software, architecture, and algorithms for supercomputing systems.

XI. CONCLUSION

The U.S. national interest in the rapidly moving area of computer architecture should be as follows:

1. To ensure that all designs likely to succeed in any major way are explored actively. We must guard against technological surprise.
2. To get successful, or potentially successful, designs into the hands of large user and software development groups quickly, so as to accelerate the development of the necessary software. We must maintain system and software leadership.
3. To accelerate the transition of successful designs into commercial production, preferably by involving strong industrial groups in their development from the start.

It is quite important for U.S. activities aimed at maintaining awareness of Japanese technical developments through suitable translations be accelerated, possibly through programs undertaken by NSF or DARPA.

Finally, since the U.S. faces rapidly mounting competition in this area from both Japan and Europe, it is important that vigorously conceived programs, including those now planned by the major agencies involved, move forward without delay.

Report of the Research Briefing Panel on Information Technology in Precollege Education

Research Briefing Panel on Information Technology in Precollege Education

John Seely Brown (*Co-chairman*), Xerox Palo Alto Research Center, Palo Alto, Calif.
James Greeno (*Co-chairman*), Graduate School of Education, University of California at Berkeley, Berkeley, Calif.
Henry Becker, Center for Social Organization of Schools, Johns Hopkins University, Baltimore, Md.
Richard Burton, Xerox Palo Alto Research Center, Palo Alto, Calif.
Allan Collins, Bolt, Beranek, and Newman Laboratories, Cambridge, Mass.
Andrea diSessa, Laboratory for Computer Science, Massachusetts Institute of Technology, Cambridge, Mass.
*Harry Lasker, Interactive Training Systems, Cambridge, Mass.
Mark Lepper, Department of Psychology, Stanford University, Stanford, Calif.
Milbrey McLaughlin, Graduate School of Education, Stanford University, Stanford, Calif.
Robert Parry, Department of Chemistry, University of Utah, Salt Lake City, Utah
Henry Pollak, Bell Communications Research, Inc., Summit, N.J.
Frederick Reif, Department of Physics and School of Education, University of California at Berkeley, Berkeley, Calif.

Staff

Norman Metzger, *Staff Officer*
Susan Newman, *Rapporteur*, The Learning Lab, WNET-TV Education Division
Gerry Kasarda, *Secretary*

*Dr. Lasker was unable to attend the meeting of the briefing panel, but met separately with the co-chairmen.

Additional advice was obtained from: Alan Lesgold, Lillian McDermott, Arthur Melmed, Andrew Molnar, Jack Palmer, Roy Pea, Senta Raizen, Paul Resta, Edward Silver, Charles Smith, John Tukey, and Fred Weingarten.

Report of the Research Briefing Panel on Information Technology in Precollege Education

SUMMARY

The development of more effective research methodologies and tools for the study of human learning, as well as recent advances in hardware capabilities, are opening up the possibility of using powerful new information technologies to aid individual learners and to help solve pervasive educational problems. Parallel work in artificial intelligence and the cognitive sciences has set the stage for qualitatively new applications of technology to education. What is required to move forward is increased support for basic interdisciplinary research, focused by the development of advanced learning systems employing the methodologies and equipment of artificial intelligence.

While success in this arena will not solve all educational problems, it will provide a new scientific basis for instructional and systems design, teacher training, and curricular restructuring. It also will create valuable new resources in the form of model electronic learning environments, while attracting a new cadre of professionals to the fields of education and educational research.

Thus, the panel recommends two federal initiatives to improve precollege education through the use of technology: (1) Support increased basic cognitive research on learning undertaken in conjunction with the development of advanced electronic learning environments employing the methodologies and equipment of artificial intelligence. (2) Continue federal support for individual scientists and applied research and development centers to analyze specific educational problems, develop and evaluate educational software, and find effective ways to integrate educational technologies into classrooms.

INTRODUCTION

Numerous recent reports have cited problems in the schools, declines in student performance on standardized tests, and the weakness of American mathematics and science education relative to other industrialized countries. Diverse sectors of society have begun to call for increased educational effectiveness and, in particular, for exploring the potential of information technologies for improving education.

However, earlier promises of educational improvements via technological interventions, such as educational television and computer-assisted instruction, have not always been fulfilled. Current investments in microcomputers are seen as limited in their effectiveness. Therefore, one charge of the present panel has been to assess what is different now—technologically, scientifically, and sociologically—that warrants increased federal attention to the research and development of instructional systems for the improvement of precollege education and what is required to advance the potential for learning represented by information technologies.

WHAT IS DIFFERENT?

While past projects to reform education based on learning research and technological advances have not always fulfilled the hopes of their designers, they have resulted in limited advances. Moreover, they have provided opportunities to learn how to achieve needed change more effectively. Recent studies of past reform efforts have identified specific difficulties surrounding school change and recommended how such difficulties can be surmounted. In addition, a confluence of several other factors is creating a unique opportunity to enhance the value of information technologies in precollege education.

Need for Educational Change

While observers are alarmed by evidence of declining "basic" mathematics and literacy skills among American students, there is simultaneously a growing emphasis on teaching "higher-order thinking skills"—reasoning, critical thinking, and learning about learning. These objectives require increased pedagogical emphasis on integrating basic skills with conceptual understanding, developing critical thinking and reading skills, enhancing the ability to see connections among objects and processes, and developing the capacity for self-directed learning.

Rapid Dissemination of Computers

The growing ubiquity of computers, while having at present relatively little direct educational impact, is nevertheless familiarizing a generation of teachers, parents, students, schools of education, school boards, and others with computers as tools and media. This sort of "computer literacy" sets the stage for the acceptance of more advanced educational applications such as those outlined in this report.

In addition, dramatic increases will occur in the computational power available to consumers at affordable prices within the next 5 to 10 years. These new levels of machine power make possible the delivery of the kinds of integrated and knowledge-based systems that are foreshadowed by present research and development efforts but that require concentrated basic research and experimentation for their full realization.

Effect of Technological Development on Research

Developments in the cognitive sciences and advances in computer technology have had a deep impact on educational research. Prior to 1960, learning research was concerned with primarily discovering the external conditions that influence readily observable behavior. While providing useful indications of the limits of human cognitive capacities and a scientific basis for the field of instructional design, behavioral psychology yielded little insight into the "black box": the underlying cognitive events that mediate behavior, including learning.

However, during the last 25 years, researchers have gained a new language and set of tools for expressing and validating learning theories that focus specifically on internal cognitive processes and their relation to the growth of knowledge and mental operations in humans. For example, symbolic computing languages used by the artificial intelligence community, such as LISP, have provided a new notational system that is both scientifically precise and sufficiently flexible to represent complex learning processes.

In addition, computer technology has brought important changes in educational research by acting as a finely tuned "cognitive lab" for testing hypotheses and manipulating the large number of variables affecting human cognition. Through the simulation of human learning and problem solving, researchers have been able to verify detailed hypotheses about the mechanisms, architecture, and language of the human mind, particularly as they affect highly skilled performance by experts or the mental work by which novices gain new skills.

Finally, owing to the rapid growth of the information technology industry, a new generation of college and graduate students is becoming well versed in the computer sciences. The result is a larger pool of scientifically trained individuals from which, given adequate incentives, it is now possible to draw talent capable of conducting the combined educational and technological research needed to address new educational objectives.

GOALS AND OPPORTUNITIES

In addition to traditional competencies, students increasingly should learn the thinking skills to manage information, formulate effective probing questions, test hypotheses, make judgments, express themselves logically and lucidly, and solve problems.

Unfortunately, our present educational resources are inadequate to meet these objectives. In most classrooms, there is a lack of integration between low-level skills and higher-level understanding of a subject, and between disciplines such as mathematics and science, or writing and other subjects. Cognitive research confirms that knowledge and skills learned without conceptual understanding or functional application to problems are either forgotten or remain inert when needed in situations that differ from those in which they were acquired.

This common dissociation of formal schooling from experiential learning prevents students from mobilizing for formal learning the power of their intuitive understanding, even though they use it effectively outside the classroom. Students receive feedback in the form of grades, but are given little opportunity to revise their school work or to reflect on or communicate with other students or the teacher about their own writing, problem-solving, or other learning efforts. As a result, many students fail to correct even minor misunderstandings that nevertheless prevent continued advancement, and to develop skills for assessing their own learning problems.

While many of the needed changes could be effected by human teachers, appropriately trained, paid, and teaching smaller and fewer classes, the costs of this approach appear to be prohibitive. The possibility exists, however, that information technologies can provide the means to deliver some aspects of instruction not only cost-effectively but also in a deeply individualized and integrated way.

Specifically, parallel research in cognitive science and artificial intelligence is within reach of developing intelligent diagnostic, coaching, and discovery learning systems for a variety of subjects. Such systems could have a major impact on education. For example, knowledge-based simulation systems can enable students to test their understanding of scientific and mathematical principles, posing original problems that reflect their own questions about the world. With more powerful computers and research, it will be possible to create sophisticated work stations that can be used over a student's school lifetime. Such systems might include computerized lab-simulators, data-analysis software, writing and idea-structuring tools, and problem-analysis and management software.

In short, the real promise for the future is the development of integrated systems capable of interacting intelligently and sensitively with students and of providing a responsive, structured environment in which various skills and concepts can be developed and practiced. To realize this potential, initial investments must be made to capitalize research equipment and to support basic artificial intelligence and cognitive research focused on education and the development of educational technology. Such investments will enable us to move beyond the usual ad hoc applications of technology to educational problems. Besides providing a better basis for curriculum design, organization of subjects, and teacher training, increased scientific research will enable us to tailor instruction for individuals to a degree not generally found outside the apprenticeship learning models prevalent in graduate schools or other research centers.

TECHNOLOGY IN PRECOLLEGE EDUCATION

Information technology applied to education can be divided into three levels: systems of hardware and software now commonly used in schools; systems currently in the development and commercial pipelines; and advanced experimental systems embodying cognitive science and principles of artificial intelligence.

Level 1: Common Uses of Technology

Most of the computers presently available in schools are 8-bit microcomputers, with limited graphics, memory, and processing capabilities; however, some elementary and high schools are purchasing 16-bit and 32-bit "personal computers." While a growing range of applications is being developed, the majority of the schools that have microcomputers now use them to teach "computer literacy," BASIC programming, and for drill-and-practice or other forms of programmed instruction.

The best versions of the latter are based on behavioral analyses of the components of skill required to perform a given learning task, together with research on simple motivational factors and reinforcement schedules that influence persistence in a task. These types of programmed instruction are those most easily implemented on computers. In addition, these systems provide students with the advantages of sequenced practice, with the problems and text material adjusted to the level of student performance. Gains in student performance of routine skills, such as arithmetic calculation and elementary programming, have been accomplished through the use of these systems.

Assessment

While providing efficient means of acquiring rote skills, the educational benefits to be gained from these baseline resources are likely to be quite modest. The limited processing power and memory of microcomputers preclude their use for learning activities that conceptual understanding with mastery of procedural skills more effectively integrate; further, the analyses on which the programs are based do not address the cognitive mechanisms by which this sort of integration can be achieved. Computer literacy courses are rapidly outmoded as new hardware and software are developed. In addition, a new generation of more powerful and flexible programming languages and techniques is being created. As a result, instruction in BASIC may provide a poor model for students who might wish eventually to tackle complex programming problems.

Level 2: Recent Innovations

More imaginative educational uses of computers are being developed commercially or in cooperation with university-based researchers. These applications, which are being used in homes and experimentally in a few schools, include simple simulations, educational games aimed at reasoning and deductive skills, and simple word-processing programs. Some of them are based on cognitive research that specifies learning problems that students have in a subject or the underlying skills that are needed for mastery of complex processes like writing. Others are based simply on the intuitions of skillful practitioners and instructional designers.

An example of a level 2 application is a simulation system, DynaTurtle, developed at the Massachusetts Institute of Technology. This system uses an extension of the programming language LOGO, in which students can give instructions that cause objects on the screen to move in accordance with the principles of Newtonian physics. Development of the system was based on the empirical recognition that students often fail to understand the fundamental concepts of a subject, even when they are able to apply formulas mechanically or to recite factual information. Cognitive research, some done using the DynaTurtle system, has documented the fact that it is often students' preconceptions about a subject that interfere with the acquisition of more accurate conceptual understandings, in spite of considerable formal instruction.

For example, students typically believe that an object in motion has some force that was transmitted to it by whatever caused the motion and that is dissipated as the motion continues. This experience-based "mental model" prevents most students from understanding momentum. Systems such as DynaTurtle take advantage of dynamic computer graphics to provide concrete information about general principles in response to student actions, enabling them to see the consequences of their ideas and to develop experience-based understandings of abstract concepts.

Assessment

Level 2 applications are being developed by educational researchers, often in collaboration with teachers and instructional designers, and frequently implemented by commercial organizations. While these instructional systems represent significant advances in educational software design, there are a number of ways in which they do not tap the full potential of information technologies to improve education:

- The systems often address isolated components of the curriculum and isolated instructional goals, and are thus difficult for teachers to integrate into classroom activities. Further, there are many learning topics for which an adequate cognitive analysis has not been performed yet, and little is known about how such specialized systems affect students' understanding of the instruction that is normally given on the same topics. Continued applied research is needed to address these concerns.
- Many of the systems have goals that are not thoroughly understood by teachers. Games addressing general cognitive skills are likely to puzzle those teachers who have a background in programmed instruction. A system for correcting a misconception in physics will seem mysterious to a teacher who shares the misconception, as many do.
- The systems that have been developed are based on ideas about helping students learn, rather than helping teachers teach. The orientation is a reasonable one, but is not optimal for encouraging enthusiastic adoption by teachers. Research and development focused on the use of computer-based materials in relation to the performance of teaching, either as a direct means of achieving teachers' goals or as supplementary materials that are viewed as beneficial, could increase the effective use of educational technology.

Level 2 applications represent a diversity of efforts that is important for discovering ways to use technology effectively for education. Besides their instructional value, such systems provide important artifacts around which to build continued research into the specific learning mechanisms that enable them to be effective. Most importantly, the involvement of teachers and instructional designers in this development makes it easier to begin implementing technological and other changes in the schools in an orderly way.

LEVEL 3: ADVANCED THEORY AND SYSTEMS DEVELOPMENT

Advanced systems combining symbolic computational capabilities, artificial intelligence, and cognitive theory have begun to be developed in a few well-equipped laboratories. As with level 2, each system addresses a relatively specific aspect of skill or knowledge but the systems differ in that they are informed by an explicit model of human learning and expertise. In addition to their eventual practical value as components of integrated learning systems, these systems and the computational equipment on which they are developed provide a powerful set of learning laboratories around which to deepen our knowledge of how conceptual change and understanding occur and of the methods by which we can bring them about more reliably.

Building on groundwork by behavioral scientists and interdisciplinary work in linguistics, psychology, and computer science, developments in cognitive science have coalesced around a theoretical model of the human mind as a complex information-processing system in which information is organized, stored, and operated on in ways that we are only now beginning to characterize in detail. Within this framework, learning may be viewed as a process of information transfer, leading to useful lines of inquiry concerning the structure of information to be transferred, the characteristics of the receiver (including the biases and constraints of the "system architecture"), and a theory of the transfer process itself.

At the same time, methods in artificial intelligence for constructing expert systems have developed rapidly. Two features of this work are important for education. First, expert systems in a given subject are useful components of computerized tutoring systems; their commercial development, therefore, represents a potentially valuable resource for education. More importantly, the construction of an expert system requires an explicit representation of the knowledge that constitutes expertise. This work, along with empirical studies of human expert performance, has been a means of gaining crucial insights into the nature of human knowledge, skilled problemsolving, and reasoning.

Expert Tutoring Systems

Early efforts to adapt expert systems to education foundered because the organization of factual and procedural knowledge within the systems bore little resemblance to human thinking and reasoning. Further work informed by cognitive research has succeeded in specifying more completely the types of knowledge required for mastery of problemsolving in a given subject, and the results have been applied to education and training. For example, developers of NEOMYCIN refined, restructured, and semantically rationalized the knowledge base of the seminal expert system MYCIN to teach medical diagnosis to students.

The panel believes that continued efforts to formalize the conceptual knowledge, problemsolving strategies, and inferential processes of experts will enable the creation of more robust systems capable of solving novel problems in a given topic or domain. Such systems should be able to tackle unanticipated, student-posed problems and, perhaps more importantly, to show students the underlying processes by which the systems choose among alternative solutions.

Expert Diagnostic Systems

Detailed specifications of the knowledge and skills required for work in a discipline make possible the construction of expert diagnostic systems that are able to isolate underlying causes of student errors. Construction of such systems requires a complete theory and symbolic representation of all the components of an overall skill. However, it also requires (a) a theory of how those subskills are mislearned that accounts for observed student mistakes; (b) a theory accounting for student modification—e.g., shortcuts—of formally taught procedures; and (c) a model of the "noise" in the students' execution of the skill.

One such system is DEBUGGY, an expert system for diagnosing procedural errors or "bugs" in base-10 subtraction. Starting from its data base of more than 130 empirically determined "primitive" or single bugs, DEBUGGY synthesizes a procedural model of the exact nature and cause of a student's deviation from a correct subtraction procedure. What makes the development of such a system especially noteworthy for education is its robustness: its ability to handle compounded student errors that arise when students have more than one bug, as well as their careless mistakes that can disguise systematic errors. Tested with over 4,000 students, DEBUGGY is able, using a set of search heuristics, to distinguish among 10^8 possible hypotheses about the nature of a student's errors.

Besides possible direct use in the classroom, computer-based diagnosis is an important part of developing cognitively based testing methods that measure a student's partial knowledge and pinpoint his misconceptions. The information provided by DEBUGGY can be used to provide personalized and precise remediation not hitherto possible.

Such systems also have been adapted for use in training student teachers, focusing their attention on the underlying causes of student learning problems and on strategies that they themselves can use in diagnosing students.

Perhaps most importantly, diagnostic systems provide a new tool for validating precise theories of learning. Previously unobserved knowledge states can now be deduced, enabling learning theorists to postulate changes in mental states and to verify them on the basis of successful prediction about observable behavior.

Expert Coaching Systems

The successful construction of expert systems and diagnostic systems in various domains has enabled the construction of computer-based coaches. Existing computer-based coaches such as WEST and WUMPUS construct a differential model of the student's conceptual strengths and weaknesses both through error diagnosis and by comparing student behavior with what the computerized domain expert would do at any given point in solving a problem. However, effective coaching also requires knowledge about how and when to intervene. Although such knowledge is usually held tacitly, even by master teachers, existing computer-based coaching systems have begun to characterize this kind of knowledge as a collection of rules that govern computer intervention.

Future work will be directed toward meeting two major challenges. The first one is to expand the current existence proofs for expert coaching systems to include a wider variety of domains, more flexible and psychologically informed tutoring rules, and more robust domain expert systems that mimic closely the reasoning processes of human experts. Such systems will provide a computational workbench for refining our understanding of tutoring knowledge and a building block for commercial educational software developers.

Secondly, systems must be both cost-effective to construct and readily modifiable by end-users. The emergence in artificial intelligence of new programming tools makes it possible to reduce the production time for a domain-specific coach from years to months. In addition, use of these sophisticated programming environments can enable researchers to make the system's tutoring knowledge more accessible and modifiable by teachers and educational technologists.

Discovery Learning Environments

Systems are also under development employing artificial intelligence techniques and based on cognitive research that does not require direct computer intervention into the learning process. Instead, these systems provide a rich environment that aids students in discovering how to solve problems more effectively by recording, structuring, and "playing back" the student's own problem solving processes.

Such systems are based on cognitive research that shows that successful learners exhibit a high degree of reflective skill, reviewing their progress toward a goal, analyzing their misunderstandings, and revising their strategies accordingly. Poorer learners tend simply to skip over or invent temporary solutions to aspects of a problem that they do not understand.

Computational systems have been developed in algebra and geometry that handle low-level computational operations, thus enabling students to focus attention on their strategic choice as they solve problems. The systems then provide a structured "trace" of their solutions so that students can see the alternative paths that they have tried. Preliminary tests with students show that they are motivated to review their choices and to formulate hypotheses leading to more effective choices in later problems.

The capability of information technologies to record and abstract students' problem-solving processes provides unique leverage for researchers interested in the role of error in learning, while providing students with a mental "mirror" of their own thinking. Armed with these tools, researchers can determine empirically the effects of increased self-reflection and of extended work and revision on fewer problems—particularly for poorer learners.

Such learning environments are impractical with the technology currently used in schools. But technology is already available in research laboratories that could make apprenticeship and experiential learning possible on personal computers of the late 1980s, along with methodologies to develop the scientific understanding of learning needed to design the computational environments and to formulate effective learning tasks.

The technology will be here; the need is for an adequate scientific understanding of how to exploit it. The next section describes five specific research problems on which the field is ready to make significant advances.

RESEARCH NEEDS

CONCEPTUAL CHANGE

Current educational practice appears to be remarkably unsuccessful at connecting school knowledge with experiential knowledge. A major research task is to redesign the educational materials, taking into account the deeply held beliefs and conceptions that students have when they begin instruction. Characterizations, already begun, of students' naive conceptions are not enough to enable remediation. A more adequate understanding of how change in the organization and representation of knowledge occurs during learning is needed.

Students' initial conceptual framework also includes powerful conceptual categories, called "phenomenological primitives," by which they organize their understanding and learning. Growing understanding of the function and origins of these structures can improve our ability to craft educational experiences and restructure domain knowledge so that students can integrate newly acquired knowledge into the informal knowledge and reasoning that they already have, while preparing for the deeper conceptual changes necessary for correct understanding.

CHANGES IN SUBJECT MATTER CONTENT

Widespread use of computation is leading to substantial changes in both the form and the content of procedural skills that it is important to teach to students. Obvious examples arise in mathematics, where procedures for calculation that are difficult to learn can be performed routinely on inexpensive calculators. Somewhat less obviously, many procedures that are taught in mathematics are done to facilitate hand calculations. For example, many procedures for "simplification" in elementary algebra produce expressions that are simpler if one has to evaluate them on paper, but are by no means simpler conceptually or for operations with a digital computer.

In addition, a major reexamination of the curriculum is called for as a result of new opportunities for integrating the content of education both within and between subject-matter disciplines. For example, the capabilities exist for creating educational data bases that can manifest connections between parts of the curriculum. Such connections are rarely made explicit. For example, Ohm's law is usually taught in elementary physics without reference to its being an example of a simple proportional function in mathematics. Design of curriculum sequences with relations like this in mind could increase significantly the coherence of the educational experience and the knowledge that results from it. Cross-referenced data bases also would aid teachers, who could generate easily examples of concepts in their instruction that relate to other areas of their students' work.

MEASUREMENT AND EVALUATION

A critical goal of research on learning is the development and testing of methods for measuring deep cognitive outcomes to guide the design of instructional systems. Current diagnostic techniques enable us to isolate misunderstandings in the acquisition of procedural skills at a new level of detail; however, the way in which more general misunderstandings develop and how they affect subsequent learning is less well understood. Research is needed to assess students' grasp of the conceptual bases and purposes of these procedural skills that they learn to perform, and to develop tests of their ability to modify a procedure flexibly when required.

The ability of sophisticated computational systems to record and abstract the actions of a student while engaging in various complex activities provides a rich source of data for learning researchers. The analysis of verbal protocols—that is, stream-of-consciousness records of subjects' thinking while performing complex cognitive tasks—has led to many advances in cognitive science. With information technologies enabling the automatic collection of electronic protocols and providing inferencing tools to aid in analyzing student thought processes, we are likely to achieve a new set of advances.

MOTIVATION

In addition to cognitive factors, intrinsic motivation is important to successful learning. One counter-intuitive result of recent research indicates that extrinsic reward—for example, an extra play period—can reduce the motivation of students who already engage in a particular activity willingly.

While motivation theory has traditionally been a "messy" and difficult area of scientific inquiry, the use of electronic learning environments as motivational laboratories, in which variables and effects can be isolated and manipulated subtly, has had a significant impact on research already. For example, study of the motivational factors of computer games—level of challenge and user control, completeness of information and its effect on curiosity, and the role of fantasy—establishes a beginning taxonomy of variables that can be examined within the context of experimentally designed learning environments. The popularity of fast-action video games and their possible consequences for the cognitive development of American youth necessitates a better understanding of how the slower, more difficult mastery of cognitive skills can be made rewarding for a broad range of students.

Research is needed also to assess the implications of what appear to be the motivating effects of the computer as a medium, independent of instructional content, and to determine its impact on the transfer of motivation and learning to nonelectronic settings. Work also is needed to determine the interactions between factors such as aptitude and a given motivational treatment within the electronic learning context and to experiment with feedback and reward structures.

Implementation Issues

As stated earlier, many past efforts to reform education have had only limited success. Analyses of educational change have identified a set of basic pedagogical techniques and curricular contents that have been particularly resistant to change. Cognitive analysis suggests that, in many cases, these are precisely what must change if improvements in the quality of learning for all students are to take place. How to achieve these changes is therefore a crucial research issue.

Appropriate sorts of demonstration projects can help to spread awareness among educators about the types of learning activities, motivational effects, and other qualitative outcomes that are possible with flexible electronic learning systems. Work at centers focusing on the use of educational technology in schools also should contribute to our knowledge about how to effect institutional change.

Curricular and pedagogical change in the direction of individualized discovery learning and greater integration among subjects are impeded by a lack of objective measures of the benefits of those forms of instruction. With the development of validatable measurements for deep cognitive outcomes and demonstrations through testing of the positive effects of cognitively based electronic learning, it may be possible to generate a discerning market demand for fundamentally new kinds of learning environments. Such an approach to needed reform has the clear advantage of mobilizing market forces, creating an accountability structure for the educational software industry, and preserving local choice in formulating school policy.

New learning systems themselves can become the carriers of cognitive theory and new principles of pedagogy into classrooms and homes. Teachers will then be able to assume a freer role in aiding and assisting students in their individual efforts. Further, teachers can use technology to improve their teaching techniques. Cognitively based learning systems can be used simultaneously to teach professional courses in education, to engage teachers in the same kinds of activity-based learning experiences that they will use with students later, and as objects of study about underlying cognitive processes.

Information technologies can have a profound impact on the sociology of the classroom. Several pilot studies have shown that the use of computers as tools and as communications devices tends to alter relationships between students and teachers and among peers. The design of collaborative activities around learning systems can help engineer a new classroom infrastructure within which innovations are less disruptive.

CONCLUSIONS

The time is ripe to develop a new scientific discipline as a fundamentally new resource for American education. A multiyear effort focused on advanced basic research using information technology to enhance and study learning could provide a major synthesis of the knowledge, theory, and methodology that now exist in the relatively separate disciplines of artificial intelligence, cognitive psychology, linguistics, and educational research in school subjects. Integral to this research synthesis will be experimentation with and construction of advanced computer-based learning environments. Such environments will serve as laboratories for deep cognitive research, while also providing significantly new educational resources directed at real educational problems.

By developing new learning environments and evaluating their pedagogical effectiveness, gains will be made in understanding the mechanisms of experiential learning, the integration of formal and informal knowledge, the integration of conceptual knowledge and procedural skill in a domain, and the integration of knowledge across present subject matter disciplines. The technological and theoretical products of this work will provide models of learning systems that can increase the number of American students who succeed in acquiring the knowledge and cognitive skills necessary for effective work and citizenship.

RESOURCES REQUIRED

In order to make these advances, increased resources of three kinds are needed, beyond the levels that are currently available in the field:

1. *Sophisticated Symbolic Computation*

Wider access to sophisticated computational equipment of the kind now used in standard artificial intelligence research is required for effective contributions to the research and development effort. At present, the availability of such equipment to cognitive and educational researchers is extremely limited, and well beyond funding allocations in these areas. The computational resources needed for cost-effective research and development exceed those required for the delivery of educational systems in schools, if they are to support the interactive programming methodology that makes possible rapid exploratory and experimental software development. Therefore, we recommend a substantial increment of funding to make sophisticated interactive computation and associated networks much more widely available than is now the case.

2. *Interdisciplinary Research*

Research on learning and on the development of educational systems development requires the talents of computer scientists, cognitive scientists, artificial intelligence researchers, social scientists, and subject matter experts. At present, opportunities for these sorts of interdisciplinary partnerships are limited and occur primarily as isolated projects in universities. A systematic effort to establish an integrated discipline, composed of individuals trained in the relatively new theoretical and technical skills needed to contribute effectively to the growth of knowledge about learning and its applications, is necessary. The creation of one or two special research centers (perhaps modeled after the Institute for Theoretical Physics at Santa Barbara—and similarly established with support by the National Science Foundation or other federal agencies) would provide powerful interdisciplinary resources to further the timely growth of the scientific knowledge requisite to significant educational applications of advanced technology.

3. *Stable Funding and Talent*

Existing programs represent a useful range of perspectives for near-term research and development, and it seems appropriate to continue with that mixture of support. However, current funding levels reflect a very low priority for basic research. Recent oscillations in funding levels at the National Science Foundation (Science Education Directorate) and at the National Institute of Education have been catastrophic. Stable funding for basic research and training is essential if this field is to attract first-rate scientists and graduate students. Programs are needed to support postdoctoral fellowships in which individuals trained in one of the relevant disciplines could visit for extended periods research institutions where they could obtain training and experience in the other fields.

General Comments

The levels of computer hardware and instructional programs now prevalent in schools, while probably beneficial, do not add sufficient value to ordinary instruction to warrant increased federal investment to promote the distribution. The more fruitful use of federal funds for distribution of technology is to continue to encourage development and integration into schools of level 2 and level 3 systems. The existing programs are needed to identify specific instructional difficulties and opportunities for improvement and to increase the use of instructional systems in schools and other educational settings.

The advanced research and development efforts suggested here are intended to complement programs that are already in progress or planned at federal agencies. Certainly, existing efforts will be enhanced significantly by the increased understanding of learning, epistemology, and advanced symbolic computation that would result from more advanced projects.

We are, therefore, suggesting a clear recognition of (1) the need to capitalize sophisticated symbolic computation equipment for use both as research tools and as hosts for experimental learning environments; (2) the interdisciplinary nature of basic research in this area; and (3) the need to attract first-rate scientists and graduate students to the field. Significant increments in research capabilities can be achieved with relatively modest amounts of additional funding for equipment and to support interdisciplinary research and postdoctoral fellowships. The results could bring about major advances in our ability to use information technologies in precollege education.

Report of the Research Briefing Panel on Chemical and Process Engineering for Biotechnology

Research Briefing Panel on Chemical and Process Engineering for Biotechnology

Arthur E. Humphrey (*Chairman*), Lehigh University, Bethlehem, Pa.
James E. Bailey, California Institute of Technology, Pasadena, Calif.
Harvey W. Blanch, University of California at Berkeley, Berkeley, Calif.
Stuart E. Builder, Genentech, Inc., San Francisco, Calif.
Robert R. Dorsch, E. I. du Pont de Nemours & Company, Inc., Wilmington, Del.
Wolf Hanisch, Cetus Corporation, Emeryville, Calif.
James Lago, Merck & Company, Inc., Rahway, N.J.
Alan S. Michaels, Alan Sherman Michaels, Sc.D., Inc., New York, N.Y.
John A. Quinn, University of Pennsylvania, Philadelphia, Pa.
Kevin Ulmer, Genex Corporation, Rockville, Md.
Daniel I. C. Wang, Massachusetts Institute of Technology, Cambridge, Mass.
Neal R. Amundson (*Ex-Officio*), University of Houston, Houston, Tex.

Staff

Robert M. Simon, *Project Director*
Robert M. Joyce, *Rapporteur*
Robin Illsley, *Administrative Secretary*

Report of the Research Briefing Panel on Chemical and Process Engineering for Biotechnology

EXECUTIVE SUMMARY

The phenomenal progress in molecular biology, genetics, and biochemistry over the past two decades now makes possible the programming of living cells of virtually every type—microbial, plant, and animal—to generate products ranging from simple molecules to complex proteins. The United States leads the world in research in this "new" biology. Maintaining that leadership and commercializing the fruits of that research are vitally important to the future economic well-being of the nation and to its international technological stature. A tremendous number of opportunities await commercial exploitation. This will not occur rapidly without adequate knowledge of the fundamental engineering principles underlying bioprocess scale-up. What is needed is a knowledge base in process engineering that combines the skills of the biologist and the chemical engineer in order to bring these opportunities to fruition. The rapid establishment of this knowledge base will require the cooperation of academia, industry, and government.

The intense international competition in commercializing high-technology opportunities, particularly in biotechnology, highlights the need for the United States to maintain its scientific and industrial leadership. Despite this nation's preeminent position in basic biological science, our foreign competitors have established commercial positions by practicing effective and forward-looking biochemical engineering. They can be expected to be no less aggressive as the results of U.S. research in the "new" biology diffuse abroad.

Estimates of the potential worldwide markets for biologically derived products fall in the range $40-$100 billion annually by the year 2000. First entry into these markets will be critically important, and major shares of the worldwide bioproducts market will be captured by those countries who possess the needed national capability in engineering research and personnel.

There are not enough biochemical engineers in the United States today, nor are there enough faculty to train the engineers who will be needed by the emerging biotechnology industry. Fewer than 20 U.S. Departments of Chemical Engineering have meaningful biochemical engineering programs; collectively, these are training fewer than 60 Ph.D. and M.S. graduates annually. This panel estimates that the annual need for graduate-level biochemical engineers over the next decade will average 2-3 times that number. Moreover, the training of biochemical engineers should expose them to the power of modern biological science in manipulating and controlling cellular biosynthesis, and to the utility of life sciences techniques and methods in solving critical large-scale bioprocess problems. Prompt and effective exploitation of the "new" biology is critically dependent on the improvement of this life science/biochemical engineering interface.

The technology needed to capitalize on the discoveries in modern biology falls in two areas:

- Increasing our biochemical engineering expertise in the design, scale-up, and optimal control of large-scale processes and culture of bacterial, plant, and animal cells.
- Expanding our knowledge base so that we will be able to achieve the large-scale recovery and purification of complex, unstable biological macromolecules, as well as of simple biologically derived organic compounds (e.g., ethanol, amino acids) from the dilute and impure solutions that are the initial products of biosynthesis.

Our recommendations for government action are aimed at developing the needed knowledge base in biotechnology and at reducing the communications gap between life scientists and biochemical engineers. Specifically, we recommend that appropriate federal agencies:

- Provide funding at adequate levels for basic research in biochemical and process engineering.
- Establish grants for cooperative, cross-disciplinary research that will stimulate the development of research groups led by joint principal investigators from both the life sciences and biochemical engineering.
- Provide funds for academic units to acquire, operate, and maintain the special types of equipment needed for biochemical engineering research.
- Establish faculty development awards to allow tenured faculty to broaden their expertise in biotechnology through special study and leaves, and to provide promising young faculty with the opportunity to immerse themselves in biotechnology research by temporarily relieving them of some of their teaching commitments.
- Establish institutional grants to train graduate students, thus providing the personnel that will be needed if U.S. industry is to capitalize on emerging opportunities in biotechnology.

INTRODUCTION

The phenomenal progress in molecular biology, genetics, and biochemistry over the past two decades now makes possible the programming of living cells of virtually every type—microbial, plant, and animal—to generate products ranging from simple molecules to complex proteins. The United States leads the world in research in this "new" biology. Maintaining that leadership and commercializing the fruits of that research are vitally important to the future economic well-being of the nation and to its international technological stature. The utilization of living cells and their enzymes for the industrial manufacture of useful products—for which the term "biotechnology" has been coined—requires the adaptation of these systems to large-scale production facilities. This transformation from laboratory experimentation to commercial production is the province of the chemical engineer, or more particularly the biochemical engineer.

Biochemical engineering translates the knowledge developed in biological and chemical science into practical industrial processing; the biochemical engineer bears the same relationship to the life scientist that the chemical process engineer does to the chemist. Chemical engineering enriched by active involvement in the "new" biology can provide a knowledge base to meet the challenges that will be faced by the next generation of chemical engineers in commercializing biotechnology. The establishment of this knowledge base will require the cooperation of academia, government, and industry. Chemistry and chemical engineering have enjoyed a long tradition of active cooperation to the benefit of the chemical industry. However, the interface between chemical engineering and biology needed for practical implementation of biotechnology has not yet evolved. What is needed is a sound curriculum that will promote this cooperation and train the biochemical engineers who can develop the knowledge base required to serve the biotechnology industry.

OPPORTUNITY ASSESSMENT

The opportunities for commercialization of the products emerging from the "new" biology are both tantalizing and highly diverse; among the more exciting prospects are applications falling into the following five categories:

Human and Animal Health Care. A revolutionary new family of diagnostic products based on enzymes, monoclonal antibodies, and other genetically engineered proteins promises to provide quick and highly accurate detection of immunity to or infection by viral and bacterial diseases, of susceptibility to autoimmune diseases, of the presence of genetic defects, or of the existence of neoplasms. Other significant opportunities include novel prophylactic products, exemplified by vaccines for the prevention of viral, bacterial, and protozoal diseases such as hepatitis, typhus, and malaria; new therapeutic biologicals for the treatment of cardiovascular and cerebrovascular disease, neurological and CNS diseases, rheumatoid arthritis, diabetes, and cancer; peptide hormonal substances that increase milk production of dairy cattle and stimulate growth, fecundity, and enhanced feed utilization of cattle and other farm animals.

Human and Animal Nutrition. Applications include the utilization of low-cost carbon sources for new microbiological and enzymatic syntheses of amino acids, sugars, and edible fats and oils, and the use of large-scale fermentation of low-value feedstocks in the manufacture of nutritionally balanced single-cell protein for human or animal consumption.

Agricultural Chemicals and Related Products. Prospects for applications of biotechnology include biologically derived fungicides and herbicides that are highly potent, highly specific, and environmentally safe; plant growth regulators to stimulate crop productivity, reduce plant sensitivity to environmental stress, and reduce fertilizer requirements; new techniques for crop propagation and strain improvement that employ genetically manipulated plant cell clones to replace seed.

Environmental Improvement and Protection. New microbial and enzymatic techniques for removing or destroying toxic pollutants in municipal and industrial wastes provide a promising horizon for the application of biotechnology to environmental problems.

Natural Resource Utilization. New microbiological processes for the recovery of metals and nonmetals (e.g., iodine) from low-grade ores and subsurface aquifers are under development, promising to increase substantially our supply of essential minerals. The possibility of increasing recovery of petroleum from depleted reservoirs by microbiological means is being explored, as are *in situ* microbial and enzymatic techniques for transforming solid fossil carbon sources (coal, shale) into gaseous and liquid fuels.

These tremendous opportunities cannot be exploited without adequate knowledge of the fundamental engineering principles underlying bioprocess scale-up. We face critical engineering problems in the design and control of large-scale cell-culture processes for the manufacture of new bioproducts, and of processes for the efficient large-scale recovery and purification of bioproducts from the complex mixtures in which they are made.

INTERNATIONAL COMPETITION IN RESEARCH

Who will lead the commercialization of the "new" biology? U.S. progress in basic biochemical engineering research has lagged substantially behind its progress in the life sciences. The United States has also lagged behind Western Europe and Japan in support for basic chemical engineering research. This situation must be corrected if our national ability to lead the commercialization of the "new" biology is not to be seriously compromised. Otherwise, the Japanese and the Europeans will certainly take the leading role. The following case histories are illustrative.

- The U.S. consumer was spared much of the world sugar crisis through bioreactor technology based on Japanese patents. This technology uses a continuous, immobilized-enzyme process that converts the glucose derived from domestic cornstarch to high-fructose corn syrup, a commercial sweetener. The process, with $1.5 billion annual sales, now uses $50 million annually of immobilized enzyme purchased from European companies.
- U.S. companies led the world in manufacture of amino acids prior to the 1960s. Then, an intensive Japanese effort began to improve amino acid production through biological and biochemical engineering techniques. First, classical genetic strain improvement techniques were applied to isolate mutant microorganisms that produce greatly increased quantities of amino acids. Second, biochemical engineers utilized very large bioreactors and developed efficient methods for recovering the product. Today, Japan dominates a business with annual world sales of $1.7 billion.
- Process technology for penicillin manufacture was developed in the United States in the 1940s. Protected by patents, several U.S. companies dominated the world penicillin market for many years. Now that the patents have expired and U.S. companies have decided to invest in research on new, higher-return products rather than on process improvement, European companies with modern process technology account for the preponderance of worldwide sales of natural penicillin.

Despite the preeminent U.S. position in basic biological science, our foreign competitors have established commercial positions by practicing effective and forward-looking biochemical engineering. They can be expected to be no less aggressive as the results of U.S. research in the "new" biology diffuse abroad. Foreign developments in biochemical engineering already attest to that aggressiveness.

- Basic technology for membrane separation of biomolecules was invented in the United States, but the Germans and Japanese are the ones applying it to separations of enzymes and amino acids from complex mixtures.
- Technology for very large (400,000 gal) continuous fermenters was developed and is being practiced in the United Kingdom. This development pushes biochemical engineering to limits not yet explored in the United States.
- Although the use of fermentation to produce ethanol is an ancient technology, more efficient immobilized-cell, continuous processes have been conceived, and Japan has established the first demonstration-scale plant.

A survey of biochemical engineering research and education in the United States, Western Europe, and Japan was carried out by several members of this panel. A striking difference between U.S. and overseas activities is the magnitude and organizational structure of the government-supported research and development effort in biotechnology. West Germany, Japan, and the United Kingdom each have three federally supported institutes dedicated to biotechnology; there are none in the United States. These nine institutes bring together academic and industrial investigators, feature cross-disciplinary activities, and have impressive operating budgets.

In 1983, the West German Gesellschaft für Biotechnologische Forschung (GBF) had an operating budget of DM37.1 million ($14 million), and the Institut für Biotechnologie II in Jülich had DM11.1 million ($4.3 million) for biochemical engineering alone. The last figure nearly equals the entire NSF annual budget in biochemical engineering. A quantitative comparison of space and associated equipment for biochemical engineering shows that West Germany's research space is nearly double that of the United States.

In Japan, the MITI 10-year plans have strong research and development components in biochemical engineering; three of the nine areas emphasized are bioreactors, animal cell culture, and membrane separation. Japanese government support of membrane separation research and development alone amounted to $20 million in 1983. The comparable NSF effort is less than $1 million.

It is obvious that countries such as Germany and Japan are laying a foundation of research and trained personnel as part of their strategy for intense international competition in the arena of biotechnology. The potential economic rewards for success are very great; estimates of the potential worldwide markets for biologically derived products fall in the range $40-$100 billion annually by the year 2000, corresponding to about 15 percent of the estimated total annual market for chemicals. First entry into these markets will be critically important in international competition, and major shares of the worldwide bioproducts market will be captured by those countries who possess the needed national capability in research and personnel.

STATUS OF U.S. RESEARCH AND EDUCATION

There are not enough technologically competent biochemical engineers in the U.S. today, nor are there enough faculty to train the engineers who will be needed by this emerging industry. According to an analysis by members of this panel, there are roughly 80 chemical engineering faculty members in the United States who currently have some involvement in biochemical engineering research and at least 50 more who have indicated an active interest in the field. Fewer than 20 U.S. Departments of Chemical Engineering have meaningful biochemical engineering programs; collectively, these are training fewer than 60 Ph.D. and M.S. graduates annually. Yet this panel estimates that the annual need for graduate-level biochemical engineers over the next decade will average 2-3 times that number. The current personnel demands of industry are recruiting younger faculty members and recent graduates who would otherwise be strengthening the university teaching and research capabilities in this rapidly growing field.

The costs of doing research in some areas of biochemical engineering are increasing rapidly as the field moves away from an Edisonian approach to an emphasis on discovering and applying fundamental engineering principles. Biochemical engineers are beginning to investigate the principles underlying the design and control of reactors for growing mammalian cell cultures. Such work holds great promise in reducing to practice the substantial U.S. lead in monoclonal antibody research and in large-scale cell culture for the production of tissue plasminogen activator for the rescue of heart attack victims. Virtually no university in the United States has cell culture laboratories equipped for engineering studies. Such laboratories are costly. A critical review of what constitutes adequate funding to perform state-of-the-art research should be part of an overall plan for ensuring U.S. competitiveness in this area.

Most important, there is a critical need to strengthen the life science/biochemical engineering interface. The historic success of the development of the U.S. chemical industry has in large measure been due to close collaboration and communication between the chemist and chemical engineer in both academia and industry. This collaboration has been fruitful because both disciplines share a common language and comprehension of physicochemical laws and principles, and a mutual appreciation for the special skills that each brings to bear on the adaptation of chemical discoveries to industrial practice. In contrast, few life scientists are aware of the engineering principles and practical problems associated with the scale-up of biological processes or with the large-scale processing of bioproducts. Conversely, few chemical engineers are sufficiently knowledgeable in the principles of modern molecular and cellular biology, microbiology, genetics, and biochemistry to permit their effective communication with life scientists. This lack of communication inhibits progress on the unique problems that make biological processes particularly difficult to engineer:

- Living organisms can undergo spontaneous mutations, which can have dramatic effects on process operations.
- Processes must be carried out under aseptic conditions to prevent contamination by ubiquitous, undesired organisms.
- Biological processes usually take place in dilute aqueous solution, imposing the need to separate the product from large amounts of water.
- Many of the high-value-in-use products of biotechnology are fragile, hard to purify, and structurally complex.

The training of biochemical engineers should expose them to the power of modern biological science in manipulating and controlling cellular biosynthesis, and to the utility of life sciences techniques and methods in solving critical large-scale bioprocess problems. Prompt and effective exploitation of the "new" biology is critically dependent on the improvement of this life science/biochemical engineering interface; this is one of the most urgent needs confronting biotechnology today.

INTELLECTUAL CHALLENGES AND CRITICAL NEEDS

Because of the rapid progress in the life sciences, commercialization of the fruits of modern biology is not opportunity-limited. The principal limitation is the lack of a technology base that can address the following key challenges:

- Increasing our existing biochemical engineering expertise in techniques of large-scale cell culture not only using simple microorganisms (e.g., bacteria and yeasts) but also utilizing the full range of plant and animal cell systems.
- Expanding our knowledge base so that we will be able to achieve large-scale processing, including the recovery and purification of complex, unstable biological macromolecules (e.g., antibodies, peptide hormones) from product mixtures, as well as the economic and efficient recovery of simple biologically derived organic compounds (e.g., ethanol, amino acids) from highly dilute and impure solutions.

Specific areas of research aimed at meeting these challenges are as follows.

BIOREACTORS

Mechanically agitated reactors currently used for antibiotic fermentations are often poorly or not at all suited to meet the diverse demands of bioprocessing. New techniques are needed for large-scale culture of plant and animal cells as well as the new microorganisms now being engineered for bioproduct synthesis. Fundamental knowledge of how physical and environmental factors influence intracellular biosynthetic pathways is essential to the development of such techniques. Bioreactor research will require the hybridization of such sciences as molecular and cellular biology, microbiology, and cell physiology with basic engineering skills, chemical kinetics, thermodynamics, fluid dynamics, heat and mass transport, and precise bioprocess control—a rare combination today in either industry or academia.

An important parallel direction of bioreactor research is the development of methods for utilizing free or immobilized enzymes, enzyme combinations, or nongrowing whole cells as catalysts for biosynthesis. The effective utilization of such diverse forms of biocatalysts requires exploration of fluid bed, fixed bed, membrane, cell recycle, tubular, and other reactor types in which to carry out biosynthesis. The challenge is to translate the existing knowledge base for chemical reactors into biosystems, where strict asepsis, complicated biological regulation processes, enzyme and cell fragility, cofactor regeneration of enzyme activity, and maintenance of cell energy provide additional dimensions to the problem. Process fundamentals developed in benchtop systems must be tested in large-scale equipment in order to gain understanding of the scale sensitivity of biosystems.

SEPARATION AND PURIFICATION

The product stream generated in a bioreactor must go through separation and purification operations in order to isolate the desired product in adequate purity. For fragile, high-unit-value products intended for human or animal health applications, there is a premium on processes that minimize product deterioration and maximize purity. For low-unit-value products where competition lies in nonbiological industrial synthesis the premium lies in recovery processes that are energy-efficient and have high recovery efficiency and low environmental impact. The areas of separation and purification science from which biotechnology can benefit through increased R&D effort fall into three principal categories:

Modification of Conventional, Large-Scale Industrial Separation Methods. Such processes as ion exchange chromatography, pressure-driven membrane separation, electrodialysis, and liquid-liquid extraction are now employed industrially to recover or purify antibotics and simple molecules such as citric acid. They are generally too nonselective or too destructive to be useful for processing fragile protein biologicals. Refinement of these techniques is needed to render them suitable for modern bioproduct recovery.

Adaptation of Biochemical Laboratory Separation Methods to Large-Scale Bioprocessing. Life scientists have developed extremely powerful and sophisticated separation tools such as electrophoretic separation and affinity separation that are potentially applicable to large-scale separations. However, much needs to be learned about the molecular mechanisms and kinetics of these processes before a rational approach to scale-up can be initiated.

Novel Separation and Purification Concepts. There is need for discovery and development of novel separation and purification concepts that embody combinations of physical, chemical, and biological phenomena outside the armamentarium of conventional chemical process technology. Several such processes are being studied in Japan and Europe. Some are based solely on physicochemical principles, for example aqueous two-phase separations employing water-soluble synthetic polymers, or multiple-field fractionation. Other process concepts represent a constructive synthesis of biochemistry and cell biology with industrial chemistry and chemical engineering. Examples include separation based on modification of permeability, separation by selective enzymatic transformation, and separation by genetically manipulated intracellular processes.

BIOPROCESS INSTRUMENTATION AND CONTROL

The successful operation of bioreactors and downstream processing equipment requires sophisticated process control, which depends upon accurate measurement of critical process variables and upon the use of advanced estimation algorithms, process models, and control strategies. Current biosensors and control methods often do not provide the desired reliability or the capability for sensitive regulation of process conditions. Furthermore, precise bioprocess control often requires on-line monitoring of the concentration of a complex biological substance for which a specific sensor does not exist. The use of enzymes, monoclonal antibodies, and even whole living cells as components of electrochemical and optical detectors offers promise in solving some of these problems.

Research on bioprocess models is needed to develop optimal control strategies and in order to extract the most useful information from measurements. Formulation of these models will require more advanced control algorithms and a greater knowledge of the effects of engineering parameters on cells and complex molecules.

TRAINED PERSONNEL

Although the need for trained personnel has already been addressed in this report, one aspect of this problem deserves mention in this section on critical needs. Constructive interaction between life scientists and engineers will be an essential element of innovative bioprocess development. The federal government can make an important contribution at this critical stage in the development of the new bioprocess science and the commercialization of biotechnology by stimulating a research environment in the academic community that will bring together life scientists and chemists with chemical engineers to focus upon bioprocessing needs and problems, and to train a generation of engineers who will be capable of designing, building, and operating the biotechnical industry of the future.

CONCLUSIONS AND RECOMMENDATIONS

If the United States is to capitalize on the revolutionary advances in biology, it must develop both a fundamental knowledge base in chemical and process engineering for biotechnology and a cadre of trained personnel so that U.S. industry can translate fundamental research results into commercial products. The infant U.S. biotechnology industry does not yet have a sufficiently broad product base to support a significant amount of fundamental research in academic institutions. In the short term, the federal government has a critical role to play in attaining the following objectives.

- Establish the fundamental knowledge base to support the design, scale-up, and optimal control of reactors and processes for the large-scale growth of microbial, plant, and animal cells required for full exploitation of the new biology.
- Develop the fundamental knowledge base for large-scale separation and purification processes that will be required to produce the spectrum of potential biochemical products from simple organic molecules to complex proteins.
- Train the next generation of biochemical engineers in a research environment that provides cross-disciplinary exchange of knowledge, industrial collaboration, availability of state-of-the-art facilities and equipment, and support levels that will assure research productivity.

The following recommendations can achieve these objectives if they are implemented at a level that will gain a competitive advantage for the United States. They will assure the development of the needed knowledge base and reduction of the communications gap between life scientists and biochemical engineers. Accordingly, we recommend that the appropriate federal agencies:

- Provide funding at adequate levels for basic research in biochemical and process engineering that will address the research areas of critical importance in this report (see Appendix).
- Establish grants for cooperative, cross-

disciplinary research that will stimulate the development of research groups led by joint principal investigators from both the life sciences and biochemical engineering.

- Provide funds for academic units to acquire, operate, and maintain the special types of equipment needed for biochemical engineering research.
- Establish faculty development awards to allow tenured faculty to broaden their expertise in the principles and techniques of biotechnology through special study and leaves, and to provide promising young faculty with the opportunity to immerse themselves in biotechnology research by relieving them from some of their teaching commitments.
- Establish institutional grants to train graduate students, thus providing the personnel that will be needed if U.S. industry is to capitalize on emerging opportunities in biotechnology.

Appendix

Research Areas of Critical Importance

A. Bioreactors

1. Quantitative Molecular and Metabolic Reaction Engineering
   - *In vivo* control mechanisms and their relationship to metabolic regulation
   - Bioenergetics of biochemical cascade reaction networks
   - Chemical regulation in biosynthesis
   - Kinetics and mechanism of action of genetically engineered microorganisms

2. Fundamental Understanding of the Physical and Environmental Factors in a Bioreactor Needed for Successful Scale-Up
   - Fluid dynamic behavior of bioreactors and its influence on biological behavior
   - Gas-liquid and liquid-solid mass transfer and heat transfer in bioreactors
   - New reactor systems for viscous fermentations
   - High-pressure, nonmechanically agitated bioreactors
   - Development of perfusion reactor systems that are compatible with large-scale production
   - Integration of bioreactor operation with product recovery
   - Bioreactors for nonaqueous systems to reduce downstream processing costs
   - Containment in bioreactors designed for genetically engineered organisms

3. Expanding the Knowledge Base on Such Biocatalysts as Immobilized Enzymes or Whole Cells
   - Mechanistic understanding of enzyme stability and inactivation
   - Techniques for immobilizing enzymes and whole cells for large-scale industrial processes
   - Bioreactor systems for reactions involving cofactor regeneration
   - Transport mechanisms in heterogeneous biocatalysis

4. Engineering Fundamental for the Design of Reactors for Handling Plant and Animal Cells
   - Quantitative definition of the stability of, and the effects of mechanical shear on, genetically engineered animal and plant cells
   - Physiological characterization of animal and plant cells
   - Characterization and fundamental understanding of surface-cell interactions for those animal cells that require attachment to surfaces in order to grow and produce ("anchorage-dependent" cells)
   - Development of surfaces for anchorage-dependent cells that can be used in large bioreactors
   - Transport phenomena in plant cells due to formation of large cell aggregates

B. Separation/Purification

1. The modification of conventional, large-scale, industrial separation methods to render them more efficient, reliable, and/or economical when employed for biological product recovery. Examples in this area include:
   - Column chromatography: development of new chromatographic materials that do not alter the structural characteristics of biological macromolecules
   - Pressure-driven membrane separation processes: membranes that are less subject to fouling by biological process streams
   - Liquid-liquid extractors: development of liquid two-phase systems that do

not alter the structural integrity of biological macromolecules

- Electrodialysis: development of membranes with selectivities useful for product streams from the "new" biology

2. Adaptation of Biochemical Laboratory Separation Methods to Large-Scale Bioprocessing

- Electrophoresis: Protein behavior in combined electrical and low gravitational fields as well as in centrifugal fields and in flow fields in gradient pore-size gels
- Affinity chromotography: mechanisms and kinetics of antibody-antigen interactions
- Affinity chromotography: factors that influence the association/dissociation equilibria in affinity separations

3. Novel Separation and Purification Process Concepts

- Fundamental principles in aqueous two-phase separations
- "Multiple Fields" such as electrical, magnetic and gravitational fractionation processes
- Separation principles and technologies based on selective enzymatic transformations
- Biological and chemical alteration of cell-wall permeability to facilitate product separation
- Integration of modern biology with downstream recovery science to alleviate potential purification problems by involving biochemical engineers in the design of recombinant cells to perform the biosynthesis

C. Bioprocess Instrumentation and Control

1. Measuring Instruments for Important Biological Variables:

- Sensors for viable cell density, biological product concentrations, cellular metabolic state, and mixed culture compositions
- More reliable and sensitive instruments for measuring concentrations of dissolved oxygen and carbon dioxide, ionic strength, and other abiotic parameters
- Biosensors based on enzymes, antibodies, and cells with adequate stability and durability

2. Mathematical Models to Aid in Process Design, Control, and Optimization:

- Process descriptions based on key features of biological mechanisms rather than on empirical correlations
- Models that accurately represent the effects of simultaneous changes in several process variables

3. Advanced Control Strategies for Regulation of Complex Biological Processes:

- Data analysis and interpretation algorithms based on biological structure and statistical filtering principles
- Control for new multiphase, multifunction bioprocesses (one example would be a bioprocess combining the reaction step with the product separation step)
- Methods for identifying and counteracting undesired cellular control mechanisms
- Controls to maximize the useful lifetime of immobilized biocatalysts and separation media

Report of the Research Briefing Panel on High-Performance Polymer Composites

Research Briefing Panel on High-Performance Polymer Composites

David McCall (*Co-Chairman*), AT&T Bell Laboratories, Murray Hill, N.J.
Rudolph Pariser (*Co-Chairman*), E. I. du Pont de Nemours & Co., Inc., Wilmington, Del.
Paul J. Flory, Stanford University, Stanford, Calif.
Frank Gerstle, Sandia Laboratories, Albuquerque, N. Mex.
Allan S. Hay, General Electric, Schenectady, N.Y.
Ted Helminiak, Wright-Patterson AFB, Ohio
Frank Kelley, University of Akron, Akron, Ohio
William J. MacKnight, University of Massachusetts, Amherst, Mass.
Arthur B. Metzner, University of Delaware, Newark, Del.
Seymour Newman, Ford Motor Company, Detroit, Mich.
Manuel Panar, E. I. du Pont de Nemours & Co., Inc., Wilmington, Del.
Darrell H. Reneker, National Bureau of Standards, Gaithersburg, Md.
L. E. Scriven, University of Minnesota, Minneapolis, Minn.
Leslie Smith, National Bureau of Standards, Gaithersburg, Md.
George M. Whitesides, Harvard University, Cambridge, Mass.
Dick J. Wilkins, General Dynamics Corporation, Fort Worth, Tex.

Staff

William Spindel, *Study Director*
Robert M. Simon, *Staff Officer*
Robert M. Joyce, *Rapporteur*
Robin Illsley, *Secretary*

Report of the Research Briefing Panel on High-Performance Polymer Composites

EXECUTIVE SUMMARY

High-performance polymer composites are formed from exceptionally strong, highly oriented, long fibers, such as graphite, aramid, or glass, which are placed in a desired orientation and bound together by a polymer matrix. The strengths and moduli thus achieved are on a par with those of the strongest available structural metals, but are substantially higher on an equal-weight basis.

The utility of advanced composites is based on their anisotropic physical properties and their unique strength and modulus per unit weight. These properties have led to significant adoptions as structural members and skins of military aircraft. As experience has been gained in the design and fabrication of these materials, they are finding increasing use in space vehicles, civilian aircraft, recreational equipment, and moving parts in textile machinery. Large-scale applications in automobiles, heavy equipment, robotics, and building and bridge construction are foreseen. We are now on the threshold of realizing the vast potential of advanced composites, which promise to play a major role in the U.S. economy and defense.

The field of advanced composites is an international arena. The present U.S. position is strong in chemistry, materials engineering and application, but Japan dominates in carbon fiber technology. In Japan, a broad MITI program covering advanced composites is in place. In West Germany, a new Max Planck Institute for polymer research will impact on composite technology.

Research needs include the following areas:

- New fiber compositions, polymer matrices, and fabrication methods are needed to broaden applications and to reduce costs.
- Understanding of composite structure/property relationships is lacking. The fiber/matrix interphase region has not been well characterized, although it is critical to composite performance.
- Knowledge of failure mechanism is primitive. Flaw identification, growth, and elimination need study. Response to differing load conditions and environmental factors has not been described in fundamental terms. Life prediction, reliability, and testing are critical matters that await characterization and understanding at the molecular level.
- A science of composite design and pro-

cessing needs to be developed and must involve extensive computer-based modeling for design, engineering, and manufacture. Joining and repair processes must also be developed on a fundamental basis. The toxicity, envionmental effects, and eventual disposal of composites must be examined for their consequences on fabriciation and use of these materials.

- Reinforcement of a polymer matrix at the molecular level, with stiff, rodlike polymer molecules, has been demonstrated and should be studied to learn how to develop optimum properties in such composites.

The rapid technological advances in high performance composites over the past decade have outpaced the development of the underlying basic science. Research in the area calls for communication and collaboration among several disciplines, including chemistry, physics, chemical engineering, mechanical engineering, and materials science. These groups tend to be insular in present U.S. university structures. The future will belong to those who can bring together the necessary disciplines and construct a coherent intellectual structure that will form the basis for advancing to new levels of understanding.

Research on advanced composites in U.S. universities is embryonic and frequently includes only one of the several science and engineering disciplines needed to advance the subject. A realignment of the traditional academic structure appears necessary if the needed interdisciplinary research is to be effective. Fewer than 30 U.S. universities conduct research on composite materials, involving about 40 equivalent full-time faculty. Multidisciplinary organizations devoted to composites exist at only two of these universities. Moreover, a substantial fraction of this university research is devoted to product design and to lower-performance short-fiber composites.

Estimates of overseas support for composites research indicate that several countries have efforts that are comparable to that of the United States. However, their programs are growing more rapidly than ours and are concentrated on the interdisciplinary approach that is essential to progress.

It is recommended that a modest number—perhaps three, as it becomes feasible to staff them—of interdisciplinary centers devoted to basic research on advanced composites be established in universities. The proposed Engineering Research Centers are a good model, and research on composites should be encouraged in them. In addition, more research specifically concerned with advanced composites should be encouraged in existing academic interdisciplinary centers devoted to materials science. The objectives of these centers would be to develop the scientific underpinning that is essential to advance the national position in advanced composites and to develop a pool of trained scientists and engineers for future personnel needs in research, development, and teaching in this high-technology area. Industrial involvement in the form of intellectual presence and research cooperation in the campus centers is regarded as essential.

INTRODUCTION

Structural composites are not new. Nature has been perfecting these materials for millions of years, resulting in such composite structures as wood, bone, or clam shells. These structures comprise combinations of components specifically chosen in terms of amount, size, shape, orientation, and interface to achieve a set of overall properties that is optimized for its intended purpose.

This is also the central purpose of manmade high-performance composites, a relatively new area of technology that holds promise for revolutionary advances in structural materials. These composite materials are filling crucial national needs. They are already essential to the defense establishment and the space program, and are beginning to have significant impact in important areas of the national economy, particularly aircraft, recreational products, textile machinery, and ground transportation.

Precursors to advanced composites are found in such diverse developments as the automobile tire and short-fiber reinforced plastics. In recent years, several classes of high-performance structural composites have been developed, including multicomponent ceramics, short-fiber reinforced metals, and polymer composites based on strong, highly oriented, continuous fibers developed through advances in inorganic and polymer chemistry. High-performance polymer composites are made by arranging arrays of strong, stiff fibers in a desired orientation, usually in close-packed alignment, and binding them together with polymer matrices. This new class of polymer composites provides combinations of strength and modulus substantially superior to those of structural metals and alloys on an equal weight basis. This report is concerned only with these high-performance polymer composites.

The principal impetus for the development of high-performance polymer composites came from military aircraft and space vehicles, which needed structural components of minimum weight and with high strength in prescribed directions. Further impetus was provided by the adoption of these composites in civilian aircraft to achieve fuel savings, in recreational equipment, and in moving parts of textile machinery to achieve higher operating speeds. We are now only on the threshold of realizing their vast potential. Significant improvements in cost and performance are expected and should greatly expand their uses.

The United States pioneered the use of advanced composites and is currently in a strong position in this technology. However, significant programs on composites research are being implemented in other countries. Specifically, a broadly based academic-industrial program, initiated by MITI, is in place in Japan. Germany's composites research will likely be strengthened soon by a new Max Planck Institute for polymer science and engineering in Mainz.

The key issue addressed in this report is the need for increased research in basic science and engineering in high-performance composites to provide the technical foundation for meeting important national needs. The rapid technological advances in high-performance composites over the past decade have greatly outpaced the development of the underlying basic science. Further improvement in such areas as composite manufacturing and processing will undoubtedly emerge from continuing applied research and development. However, basic research in this area must be strengthened and broadened to acquire needed knowledge and to assure a supply of professional scientists and engineers with the background necessary to advance the technology. An investment in basic research would be a high-leverage investment.

This report is organized into the following sections: current and projected APPLICATIONS of high-performance composites; SCIENCE AND TECHNOLOGY—the present and prospective state of the art, international competition, and goals and directions for research; EDUCATIONAL NEEDS; and CONCLUSIONS AND RECOMMENDATIONS.

APPLICATIONS

The evolution of high-performance composites is founded in the opportunity to build light-weight, high-strength, high-modulus structural parts that can be tailored to meet specific loading conditions. Because these composites can be made with superior strength and modulus in one or two directions, structural parts can be designed specifically to accommodate the anticipated direction of loading, in contrast to isotropic metallic parts. Advanced composites are currently high in cost, reflecting costs of fibers and manufacturing processes; both cost elements are decreasing and should continue to do so.

High-performance composites are widely used in space vehicles because they lower launch weight and have good dimensional stability over a wide range of temperatures owing to their low coefficients of thermal expansion. These composites are used as structural components in many military and civilian aircraft and will comprise a larger fraction of the structural members in the coming generation of aircraft. Range, payload, fuel consumption, tooling costs for parts, and radar profile all enter into design considerations, and the payoff can be enormous. Japan has identified the aircraft industry as one of five in which they expect to be a leader in the 1990s, and their commitment to composites research is part of their strategy.

Sporting goods, including spars for boats, racing car bodies, bicycle frames, tennis racquet frames, fishing rods, and golf club shafts, bring high-performance composites to the public eye. Composites provide superior performance, and users are willing to pay for it.

Use of advanced composites in automobiles and other surface vehicles has been limited, and a number of problems must be solved before this industry can realize the benefits of these materials. An acceptable balance between processing speed and product quality has not yet been achieved. Useful technology for joining and repairing is not in hand, and long-term dimensional stability must be improved. However, there are important benefits to be gained by using composites in such vehicles. Tooling costs are much lower than for steel, providing greater manufacturing flexibility, more rapid design changeover, and reduced capital investment. It has been estimated that the 1,500 body parts in one automobile could be reduced to a substantially lower number by substituting high-performance composites for metals. Composites are less subject to corrosion than metals, and vehicle weight is reduced, with consequent fuel savings. At present, advanced composites are finding applications in leaf springs and drive shafts in production vehicles, and experimental composite engines have been constructed and tested. Automotive applications in the United States could expand dramatically in response to lower composite costs and the development of lower-cost manufacturing processes for using them.

Military uses in addition to aircraft are many and growing. As in sports equipment, performance rather than cost is the driving force. Examples include helmets, body armor, parts for personnel vehicles and for tanks, light-weight segments for movable bridges, carriages and braces for artillery, and portable rocket launchers.

High-performance composites have great promise in industrial machinery as reciprocating, oscillating, or rotating parts where the weight, strength, fatigue life, noise, and natural vibrational frequency of metals limit operating speeds and thus productivity. These composites are being used for moving parts in textile machinery, where higher operating speeds and productivity are paramount. A carbon fiber/epoxy centrifuge rotor can be spun at twice the speed of one made of high-tensile steel because of its high specific strength, providing a separation efficiency four times that of the steel rotor. Significant productivity increases can be foreseen by taking advantage of the high specific strength and modulus, lower moment of inertia, and the self-damping characteristics of composite parts.

High-performance composite parts have potential uses in robots and heavy construction machinery, where the low moment of inertia of a cantilevered member translates into lower energy costs and more accurate control of its positioning. Another potentially large area of use is load-bearing members of buildings and bridges, where the strength/weight ratio, self-damping, and corrosion resistance offer substantial benefits.

In summary, advanced composites are being used in a number of industries that are vital to the national economy. Their uses in these industries will unquestionably expand with anticipated improvements in cost, manufacturing speed and consistency, and particularly in basic understanding of the critical properties that must be measured to ensure optimum design and prediction of the useful life of structural parts.

SCIENCE AND TECHNOLOGY

Fibers

High-strength fibers are the primary load-bearing components of high-performance composites; they comprise about 60 percent of the volume of the composite and largely control both properties and cost. Carbon, aramid, and glass are the most widely used fibers. They have tensile strength/density and modulus/density ratios several times that of steel and represent a new technology that is just beginning to demonstrate its potential.

Glass fibers were the basis for some of the earliest reinforced polymers. The development of composites of continuous glass fibers, although largely empirical, led to the first controlled processing conditions and microstructure that took advantage of the fiber strength.

Carbon fibers are stronger, stiffer, and more costly than glass. They are produced by pyrolyzing oriented polyacrylonitrile to leave carbon, which is graphitized. The process may include hot stretching. Lower-performance carbon fibers are made from coal tar or petroleum pitch. Empirical adjustment of processing parameters has resulted in commercial carbon fibers in which either tensile strength or modulus is optimized. The strain of carbon fibers at break is about 1 percent.

Aramid fibers are prepared by spinning a polymer that has a highly rigid linear molecular structure from an oriented anisotropic phase, in which the molecules are spontaneously oriented over microscopic dimensions. During spinning the oriented regions are aligned with the fiber axis to give high strength and rigidity. Because of the low density of the polymer (1.44), the specific tensile strength of the fiber is more than 5 times that of steel. The elongation at break is greater than that of nonmetallic fibers of similar modulus, and this property can be an advantage in composite design because toughness is related to energy absorption before break. Aramid fibers have excellent thermal properties (although inferior to those of inorganics), good chemical stability, and low flammability. Like carbon, they can be made in a range of properties that optimize either strength or modulus.

The spinning of an anisotropic phase has been extended to both solutions and melts of newer polymers. For example, experimental samples of polymers such as polybenzothiazole appear to have better properties for composite use than the commercial aramids. The spinning of anisotropic melts offers the potential of lower costs than spinning from solution, but the polymers that can be spun from melts tend to give fibers with lower use temperatures. Another example is polyethylene, which usually has low modulus and strength. It has been fabricated into ultrahigh-strength fibers by gel spinning. Processes for making fibers from poly(ethylene terephthalate) have recently been patented by a Japanese group.

Boron fibers are prepared by decomposition of a volatile boron compound on a resistively heated fine tungsten filament, which is converted into tungsten boride and serves as the fiber core. Fibers of silicon carbide can be made by a similar process or by pyrolysis of a polycarbosilane fiber. Fibers of silicon nitride and alumina have also been developed.

Since the various fibers have different mechanical properties as well as costs, it is not uncommon to create hybrid composite structures comprising more than one kind of fiber to balance performance/cost requirements.

Tensile failure in carbon, glass, and other inorganic fibers is typically initiated by a flaw

TABLE 1 Properties of Fibers

		Steel	Aluminum	E-Glass	S-Glass	High-Strength Graphite	High-Modulus Graphite	Aramid	Boron
Tensile strength (10^3 psi)	Achieved 1984	500	100	350	650	720	350	525	500
	Theoretical	3,000	1,000	1,000	1,280	6,000	6,000	4,000	5,800
Tensile modulus (10^6 psi)	Achieved 1984	30	10	10	12.8	36	75	18	58
	Theoretical	30	10	—	—	100+	100+	29	—
Density (g/cm^3)		7.8	2.7	2.54	2.5	1.75	1.8	1.44	2.7
Thermal expansion (10^{-6}/°c)		11	24	5	3	−1	−1	−2	5
Specific strength (10^6 in.)		1.8	1.0	3.9	7.3	11.5	5.4	10.5	5.2
Specific modulus (10^8 in.)		1.1	1.0	1.1	1.4	5.8	11.7	3.5	6.0

or surface growth step that causes local stress concentration and produces a simple fracture perpendicular to the fiber axis. The weak lateral bonding and oriented polycrystalline nature of organic fibers such as aramid produce a fibrillar failure. Considerable splitting along the fiber axis occurs and results in a mixed shear and tensile failure.

In compression, carbon fibers usually fail by combined interface separation and microbuckling. Aramid develops shear or kink bands that form at relatively low stresses but do not produce catastrophic failure. Glass and boron fibers fail in compression when an axially oriented flaw is activated by expansion in the noncompressed direction. Compressive strengths of boron and glass composites, as well as those made from alumina and silicon carbide, exceed their tensile values.

The origin, growth, and interaction of fiber flaws with the composite matrix are poorly understood. Failure mode is not predictable from tensile strength and modulus, and the question of what measurable properties are critical to behavior in use has not been resolved. It is recognized that the time-dependent behavior of fibers under stress is important. Most test numbers are derived from short-term static tests, which do not give a reliable indication of long-term life in use.

The strength and modulus of current carbon, aramid, and glass fibers are significantly below the calculated theoretical values. However, substantial improvements in the properties of all three have been made over the last 10 years; while it is unlikely that the theoretical properties will be attained, there is good reason to expect that improvements will continue to be made. Moreover, the possibilities in organic polymers for new high-strength, high-modulus fibers have only begun to be explored, and it is expected that organic fibers with strengths 2-3 times those currently available will be made.

MATRICES

The matrix component of a composite serves primarily to bind the load-bearing fibers. The matrix is the minor component of the composite, and, while it does not usually support stress, flaws or damage in it may lead to failure of the composite. The properties of the matrix usually determine the limits of temperature and chemical environment that the composite can withstand.

Current polymer matrix materials are principally thermosetting epoxy resins, which are chemically cured to produce highly cross-linked structures. The curing process requires elevated temperatures and substantial times. Cross-linked epoxies have very low fracture toughness, and their composites are intolerant of damage because cracks will propagate readily through the matrix. Tougher materials have been developed by incorporating a rubbery dispersed phase in the epoxy, but these polymers have lower continuous use temperatures and lower modulus. Epoxy composites can be used at temperatures up to about 175°C. However, under hot/wet conditions they lose properties at lower temperatures. New high-temperature/high-modulus resins, such as polyimides, bismaleimides, and acetylene-terminated oligomers based on high-temperature resins, promise improved matrix properties with lower moisture sensitivity and faster curing. These matrices offer the potential of use temperatures in the range 200-370°C. However, their toughness is still low, and they must also be cured for significant times at elevated temperatures.

Tough, high-temperature-resistant and solvent-resistant thermoplastics with low moisture sensitivity are being explored as matrix resins. Examples are polyamideimides, polyphenylene sulfide, polyetherimides, and polyether-etherketone (PEEK). They can be processed in the melt or in solvents that must subsequently be evaporated. They are strong candidates for replacing thermosets because of their shorter fabrication cycle, which does not require a chemical cure. Property improvements should result from their greater inherent toughness and the fact that some of them are less water-sensitive than thermosets. Thermoplastics creep under stress at elevated temperatures, and current materials have use temperatures in the range 150-230°C.

The ideal resin for a matrix material would be a low-viscosity material that could easily be combined with the reinforcing fiber and could then be converted rapidly and reproducibly at low temperatures into the final composite structure. Important properties of the matrix include high and/or low temperature capability, resistance to water and solvents, toughness, and a coefficient of thermal expansion that matches that of the fiber. Although it is unlikely that all of these properties will be found in one matrix material, the potential of new polymer-forming monomers and reactions, copolymers, block copolymers, and blends should lead to a variety of high-performance matrices.

COMPOSITES

One method of manufacturing high-performance composites involves the resin-impregnation of bundled continuous filaments of a reinforcing fiber to form a continuous tape or fabric. This prepreg is subsequently plied in layers that may have the same or different alignments. The plied structure is then subjected to pressure and heat, which yields a consolidated composite structure for subsequent assembly or manufacturing steps. The combination of structural design efficiency and the flexibility of the manufacturing process through the layering of directional plies provides the desired structural properties.

Another manufacturing process involves winding a coated fiber on to a mandrel. A computer-controlled payout system capable of multiaxis precision winding can generate virtually any geometry on an irregularly shaped mandrel. The fiber-wound structure is impregnated with polymer, which is then cured, and the composite structure removed from the mandrel.

Although some aspects of composite technology are quite advanced, in others our scientific understanding and processing methods are clearly inferior to, for example, the knowledge of structure-property relationships and fabrication techniques for steel or aluminum alloys. Since structural metals are the principal competition for composites, increased acceptance of high-performance composites depends on their becoming as reliable, simple, and economical to produce as highly developed alloys.

High-performance composites have presented engineers and materials scientists with a variety of new challenges. Analytical and numerical techniques had to be developed to treat anisotropic mechanics problems. In particular, linear elastic mechanics needed for orthotropic plates and shells, and the micromechanics for calculating homogeneous, thermomechanical properties are now in hand. Techniques also exist for predicting thermoelastic properties and for detecting mechanical or compositional flaws for the more common composite components. Processing and fabrication of thin laminates into consistently high-quality materials have become routine.

However, composite manufacturers have not yet developed techniques for consistent production of thick laminates for highly stressed structural elements; fibers often move during the fabrication process, with deterioration in composite properties. Producing a void-free, high-fiber-volume laminate still involves much empiricism, since the process of creating and curing a laminate is difficult to model. Aspects of high-performance composites that are less developed, but also needed for design, are data on dynamic tensile and fatigue strength. Progress has been made in predicting the effect of flaws on failure modes dominated by the resin, but not those dominated by the fiber.

COMPETITIVE WORLD POSITION

The United States has a strong position in aramid fibers and a reasonable level of activity in developing other organic fibers with comparable or better properties. However, aramid fibers may be soon commercialized in both Europe and Japan, and the activity on other organic fibers appears to be at least as strong in Japan. MITI has designated the development of a "third-generation" fiber as a government-subsidized project, beginning in 1983 and targeted for practical applications in 5 years. Because of the overwhelming importance of the load-bearing fibers in composites, the field is sensitive to breakthroughs in stronger fibers. Current patent activity suggests that such a breakthrough is as likely to come from Japan as from this country. For example, processes to make high-strength fibers from poly(ethylene terephthalate) have recently been patented by a Japanese group.

Much of the technology used for manufacture of carbon fibers in the United States is licensed from Japanese companies. Although there is research activity in this country, some companies have chosen joint ventures with Japanese firms to take advantage of Japan's developed technology. The high level of Japanese carbon fiber technology also suggests that they may produce many of the expected future advances in these materials.

The United States currently has a strong position in composite manufacturing and processing technology and leads the world in developing major applications in such industries as aircraft, sporting goods, and automotive. However, there is growing oversees activity in composites technology. Foreign manufacturers are gaining valuable experience by subcontracting for the manufacture of large structural parts for new-generation U.S. commercial aircraft, using technology supplied by the United States. European countries are as active as the United States in putting high-performance composite parts into their military aircraft. Japanese and West European technologies are beginning to rival ours.

RESEARCH DIRECTIONS

The rapid technological development of advanced composites has outpaced the basic understanding of the structure and performance of these materials. A strengthening of basic research in this area is needed to provide a firm scientific foundation for the anticipated future advances in composite technology. General directions for such research are summarized below.

We need better understanding of molecular structure/property relationships of fibers and of how such properties translate into their behavior in a composite. Composites are subjected to long-term cyclical loading, and we need to understand the causes of fiber fatigue under dynamic conditions. Although there is some knowledge of how different fibers fail under tensile and compressive loads, there is essentially no knowledge of how such failure may be dictated by the fiber microstructure. We also need to investigate the effects on composite properties of fiber molecular structure and macrostructure: molecular composition, molecular orientation, surface structure, size, and geometrical placement in the composite, and how composite properties are affected by processing.

We need to define the physical and chemical parameters of matrix resins that can be translated into prediction of performance in a composite. The curing of thermosetting matrices should be studied to determine the effects of curing conditions on the cross-linked network structure, physical properties, and fiber interactions with the cured resin. Structural modifications of epoxies and other thermosets that will increase their toughness should be explored. Thermoplastic polymers have potential advantages over thermosets in offering greater toughness and requiring shorter composite fabrication times. The interaction of thermoplastic polymers with fibers at the interface will be substantially different from that of thermosets and need to be understood at the molecular and morphological levels. Research may point the way to reducing the tendency of thermoplastics to creep at elevated temperatures under load.

There is much to be learned about the fundamentals of the fiber-polymer interface: How does adhesion affect the load bearing, and how can optimum adhesion be achieved; what are the molecular structure and mobility of the polymer at the interface; what can be learned about nonequilibrium polymer conformations at the interface; how does the interface affect the kinetics of growing polymer chains; what properties can be achieved with various polymers and fibers when the interface is prepared under ideal laboratory conditions?

We lack the knowledge base for combining fibers and matrices to take maximum advantage of their properties. Nondestructive test methods must be developed to determine whether the manufacturing process has attained the desired microscopic and macroscopic structures, with the fibers in their desired positions. We need more precise techniques for predicting composite properties from properties of components, for detecting and predicting effects of flaws, and for predicting remaining service life of a part in use. Because of our limited basic knowledge, composite structures are now designed with considerably greater margins of safety than those used for metal structures, thereby losing some of the inherent superiority of composites. We need better techniques for joining or fastening composite parts to metals with minumum adverse effects on composite strength.

Many potential applications for advanced composites will depend on achieving substantial improvements over current manufacturing speeds and product quality. We must develop automated robotic manufacturing processes to reduce product variability. Composite samples prepared under carefully controlled laboratory conditions should provide benchmark properties against which to assess manufactured parts.

The toxicity of composite components, their long-term environmental effects, and their recyclability need to be studied, and should be considered at the development stage of a composite.

Finally, we should draw upon knowledge of the structure of natural composites such as bone, wood, and clam shells for clues to how man-made composites can be made to approach more nearly their theoretical physical properties.

EDUCATIONAL NEEDS

The rapid technological advances in high-performance composites over the past decade have outpaced the development of the underlying basic science. Research in the area calls for communication and collaboration among several disciplines, including chemistry, physics, chemical engineering, mechanical engineering, and material science. These groups tend to be insular in present U.S. university structures. The future will belong to those who can bring together the necessary disciplines and construct a coherent intellectual structure that will form the basis for understanding present achievements and advancing to new levels.

Research on advanced composites in U.S. universities is embryonic and frequently includes only one of the several science and engineering disciplines needed to advance the subject. A realignment of the traditional academic structure appears necessary if the needed interdisciplinary research is to be effective. Fewer than 30 U.S. universities conduct research on composite materials, involving about 40 full-time equivalent faculty. Multidisciplinary organizations devoted to composites exist at only two of these universities. Moreover, a substantial fraction of this university research is devoted to product design and to lower-performance short-fiber composites. Consequently, too few scientists and engineers are being trained with the broad background needed to advance composite technology, and the gap between supply and demand for trained people is increasing with the rapid growth of the composites industry.

Estimates of overseas support for composites research indicate that several countries have efforts that are comparable to that of the United States. However, their programs are growing more rapidly than ours and appear to be concentrated on the interdisciplinary approach that is essential to progress.

CONCLUSIONS AND RECOMMENDATIONS

Our conclusions are as follows:

- High-performance composites will find increasing numbers of applications in both industry and defense. They are essential to many high-technology industries and they will meet important national needs.
- The present U.S. position in high-performance composites is strong in the areas of chemistry, materials engineering, and applications. There is a strong challenge to the United States in this area from both Japan and Europe.
- Basic research on high-performance composites in the United States must be strengthened to develop the understanding needed to advance the field and to maintain the U.S. position against foreign competition.

It is recommended that a modest number—perhaps three, as it becomes feasible to staff them—of interdisciplinary centers devoted to basic research on advanced composites be established in universities. The proposed Engineering Research Centers are a good model, and research on composites should be encouraged in them. In addition, more research specifically concerned with advanced composites should be encouraged in existing academic interdisciplinary centers devoted to materials science.

The objectives of these centers would be to develop the scientific underpinning that is essential to advance the national position in advanced composites and to develop a pool of trained scientists and engineers for future personnel needs in research, development, and teaching in this high-technology area. Industrial involvement in the form of intellectual presence and research cooperation in the campus centers is regarded as essential.

Report of the Research Briefing Panel on the Biology of Oncogenes

Research Briefing Panel on the Biology of Oncogenes

Robert A. Weinberg (*Chairman*), Associate Professor of Biology, Massachusetts Institute of Technology, Whitehead Institute, Cambridge, Mass.
Raymond Erikson, American Cancer Society Professor of Cellular and Developmental Biology, Biological Laboratories, Harvard University, Cambridge, Mass.
Fred Rapp, Professor and Chairman, Department of Microbiology, The Milton S. Hershey Medical Center, Pennsylvania State University, Hershey, Pa.
Michael Sporn, Chief, Laboratory of Chemoprevention, National Cancer Institute, National Institutes of Health, Bethesda, Md.
George Todaro, Scientific Director, Oncogene, Seattle, Wash.
George Vande Woude, Director, LBI Basic Research Program, NCI Frederick Cancer Research Facility, Frederick, Md.
Michael Wigler, Senior Staff Scientist, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Rapporteur

Luis F. Parada, Massachusetts Institute of Technology, Whitehead Institute, Cambridge, Mass.

Staff

Barbara Filner, *Director,* Division of Health Sciences Policy, Institute of Medicine
Barbara Mandula, *Consultant*
Naomi Hudson, *Administrative Secretary*

Report of the Research Briefing Panel on Biology of Oncogenes

ONCOGENES AND THE MOLECULAR BASIS OF CANCER

INTRODUCTION

Researchers have sought to understand the causes of cancer for many decades. The advances prior to the early 1970s were largely descriptive. However, over the last 10 years, new ideas have been formulated and better technologies developed to address the problem. Possibly the most important result of these labors has been the recent exciting discovery of cellular oncogenes, or cancer genes. Cancer no longer appears as a vaguely defined collection of phenomena with no central mechanistic basis. Researchers are now beginning to define in precise terms—down to the last molecule—the kinds of changes that occur within a normal cell to propel it forward to becoming a tumor cell.

Various agents from outside the organism, such as radiation, chemical substances, and certain types of viruses, have been identified as involved in cancer induction. The accumulated evidence now suggests that viruses and radiation are involved in causing only a minority of human cancers. The suggestion that chemical substances (including agents in tobacco smoke and in the diet) induce the bulk of human cancers did not itself shed light on the molecular processes involved in the transformation of normal cells into tumor cells. A central question remained: How does a cancer-causing chemical alter a cell so that this cell and its descendants exhibit cancer traits? It appeared that the genetic blueprint of these cells, carried in their DNA, must be permanently changed (mutated). This implicated DNA molecules as the central target of carcinogenesis.

The chasm between the formulation of this explanation and a demonstration of its correctness could be bridged only by recent advances in cellular and molecular biology. These advances, both conceptual and technical, led to the discovery of cellular oncogenes and some of the molecular mechanisms of carcinogenesis. Modern DNA technology, developed in studies seemingly unrelated to cancer, has been mustered to isolate specific DNA sequences, analyze them in great detail, and begin to relate their properties to the process of carcinogenesis.

THE SEARCH FOR ONCOGENES

Retrovirus-Associated Oncogenes

Our initial knowledge of cellular oncogenes stemmed from tumor virology. RNA tumor viruses ("retroviruses") were known to cause cancer in lower animals. These viruses harbor cancer-causing genes, called viral oncogenes, that are not genuine viral genes, but rather genes from animal cells that were captured by the viruses in the course of infecting cells. These acquired, previously normal, cellular genes fell under the control of the viral genetic machinery. The result was that a normal cellular gene, a "proto-oncogene," was converted into a potent oncogene. There are more than 20 different oncogenes now known to be associated with a variety of retroviruses, each pointing to the existence of an antecedent proto-oncogene residing among the genes of normal cells.

The normal human genome has precise counterparts (homologues) to the proto-oncogenes originally found in animal genomes. Proto-oncogenes also have been detected in such other organisms as yeast, fruit flies, and mice. The conservation of these proto-oncogenes during evolution suggests that they play a crucial role in cellular and organismic survival.

Discovery of Human Oncogenes

Gene transfer or "transfection" techniques make it possible to extract genes from one cell and introduce them into another. Transfection experiments showed that DNA segments from human cancer cells contain genetic information able to transform normal recipient cells into tumor cells. These DNA molecules contained specific transforming segments termed, once again, oncogenes. Such oncogenes were found in a wide variety of human tumors having no known viral etiology. All the isolated oncogenes were very similar to genes found in normal cells. This confirmed the suspicion that *oncogenes of human tumor cells arise by alteration of genes present in normal human cells.*

Subsequent results were surprising in two ways. First, the human oncogenes were closely related in structure to the oncogenes that have been acquired and activated by cancer-causing animal retroviruses. Thus, mechanisms of cancer causation in humans and animals apparently have common molecular features. By extension, all types of cellular proto-oncogenes that had been activated by animal retroviruses could potentially also be activated in human cancers.

The second surprise occurred when researchers first compared the molecular sequence of a human cancer gene, isolated by recombinant DNA technology, with that of the corresponding proto-oncogene that resides in a normal cell. They found that this oncogene, present in a human urinary bladder carcinoma, had become activated by the mutation of just one nucleotide among the 5,000 nucleotides comprising the normal proto-oncogene. This change resulted in one different amino acid in the protein encoded by that gene. Thus, a very limited change in the structure of a normal protein resulted in a new mandate for the cell that harbored it. That first evidence— now much expanded— showed that important steps in carcinogenesis can be traced to specific, well-defined genetic lesions. Presumably, the mutation responsible for the activation of this particular proto-oncogene occurred in bladder tissue during the course of tumor formation.

About 15 to 20 percent of all human tumors, from a wide variety of body sites, have been shown to carry oncogenes in their DNA. Most of these oncogenes are members of the *ras* gene family. (Oncogenes are given three-letter names, such as *ras*, *src*, *myc*, *mos*; *ras* stands for rat sarcoma, the tumor in which it was first found.) Some of the *ras* genes have been activated in experimental animals with x-radiation and with chemicals known to be mutagenic and carcinogenic. Transfection assays continue to unveil new human oncogenes, at least six of which are now being characterized in laboratories around the world.

Oncogenes and Chromosomal Abnormalities in Malignancies

Many types of human cancers carry specific chromosomal rearrangements. The role of these rearrangements in cancer development was not understood until molecular analysis of human proto-oncogenes became possible.

The best examples of specific chromosomal rearrangements are the "translocations" found in human Burkitt's lymphoma and in mouse plasmacytomas. In both of these diseases, a chromosomal rearrangement relocates the normal *myc* proto-oncogene from its usual chromosomal locus into the vicinity of antibody-producing genes. As a result of this chromosomal translocation, the previously normal *myc* gene is no longer regulated by its usual control mechanisms. With loss of regulation, the *myc* gene is converted into an active oncogene that contributes to the creation of a tumor cell.

In addition to translocations, cancer cells often carry genes in amplified copy number (many copies of the same gene). DNA analysis of some of these amplified genes shows that they are oncogenes. The N-*myc* oncogene is amplified in advanced stages of human neuroblastoma; the *ras* genes have been detected in amplified form in several human cancers; and the *myc* oncogene is amplified in promyelocytic leukemia, colon carcinomas, and small cell lung carcinomas. The amplified copies of these genes cause overproduction of gene-associated products.

BIOCHEMICAL CHARACTERIZATION OF ONCOGENE PROTEINS

Oncogenes exert their effect on cells by means of the proteins whose production is specified by the nucleotide sequences of the genes. The amino acid sequences of almost all the known oncogene proteins have been determined. In many cases, the oncogene proteins closely resemble normal cellular proteins. This close relationship suggests that the oncogene proteins function similarly to normal cellular proteins, but in a deregulated or inappropriate way.

The best characterized of the oncogene proteins is that specified by the *src* oncogene. This protein, associated with the inner surface of the cell membrane, acts as a protein kinase; it attaches phosphate groups covalently to other proteins or to itself. These phosphate groups are linked to the tyrosine amino acids of the target proteins. This specificity distinguishes oncogene kinases from other protein kinases of the cell, which attach phosphate groups to the serine or threonine moieties of their targets. Recently, the *src* protein and a related protein have been found to phosphorylate certain membrane phospholipids. Proteins specified by the *ras* oncogene group have a contrasting behavior, in that they seem to be able to bind or degrade guanosine nucleotides.

Oncogenes and Growth Factors

Recent work has shown that the gene product of one retroviral oncogene, designated *sis*, is related to a subunit of platelet-derived growth factor, a protein actively involved in tissue repair. Similarly the gene product of a second retroviral oncogene, *erb*-B, represents a fragment of the cellular receptor for epidermal growth factor. These findings not only have great import for identifying the nature and mechanism of action of certain oncogenes, but also have brought together two exciting fields of study: oncogenes and growth factors. The work suggests that the intercellular signals for regulating growth (the growth factors) are intimately connected with intracellular mechanisms of growth regulation (oncogene proteins).

MAJOR PROBLEMS IN THE FIELD

How Oncogenes Work

Having learned to detect oncogenes and their encoded proteins, researchers now are faced with the more difficult task of discovering how individual oncogenes work to cause cancer. It is assumed that the proteins encoded by oncogenes affect the normal cellular pathways that govern growth. How do they do this?

The *src* and *ras* proteins are representatives of the large class of oncogene proteins that are found in the cytoplasm of the cell or near the inner surface of the cell membrane. Other proteins, specified by the *myb* and *myc* oncogenes, are found in the nucleus, where they may control expression of certain cellular genes. The enzymatic functions of these various proteins can provide clues to their physiological roles. Of related interest are cellular regulatory processes in which simple actions of enzymes can result in large signal amplification and produce diverse responses in the cell. Presumably, each of the oncogene proteins will eventually be found to have an enzymatic or regulatory activity that allows it to perturb a cellular growth-regulating pathway. But for the moment, we are ignorant of the nature of these growth-regulating circuits and pathways and of the mechanisms by which relatively simple proteins can elicit the multiplicity of traits associated with cancer cells. This problem of oncogene protein function represents the major challenge of the field for the coming decade.

Interaction of Oncogenes with the Cell Cycle

The simplest view of the growth status of the cell would portray two states—growing and nongrowing. In reality, a complex series of events occurs when a cell emerges from a quiescent state and enters a state of exponential growth. This latter state is characterized by a well-ordered series of events, termed the "cell cycle," that includes chromosomal replication and mitosis. It is probable that some of the complex regulatory decisions required during the cell cycle involve proto-oncogenes. For example, a rapid increase in expression of the normal *myc* gene is one of the initial events that occurs during emergence from quiescence.

Cellular Differentiation

Cancer is a disorder of differentiation as well as of cell proliferation. Stem cells in many tissues form a pool of proliferating cells, some of whose progeny leave the pool and evolve (i.e., differentiate) into specialized cells that constitute the functioning tissue. Stem cells display immortality; they can, in principle, multiply without limitation throughout the lifetime of the organism. Their differentiated daughter cells, however, lose their ability to divide indefinitely. One mechanism by which cancer cells can achieve unlimited replicative potential may depend on inhibiting the ability of these cells to differentiate, trapping them in the pool of replicating stem cells.

One possible role of oncogenes is to maintain the tumor cell in an undifferentiated state, protecting it from progression to the differentiated state with concomitant loss of replicative ability. The process of differentiation and its regulation by (proto-) oncogenes represent major problems in cell biology.

Which Cancer Cell Phenotypes Are Induced by Oncogenes?

The most commonly used definition of an oncogene implies an ability, on the part of a gene, to confer unrestrained, abnormal growth capacities on previously normal cells. Such a definition would suggest that many of the genes involved in the control of normal cell proliferation have the potential, following appropriate activation, to become oncogenes.

However, cancer is more than a problem of unrestrained growth. Cancer cells acquire a number of other traits during their long progression from normality to malignancy. One important trait concerns the ability of these cells to evade being killed by components of the immune system that are responsible for detecting and eliminating abnormal cells that may arise in the body. A second attribute of some cancer cells, metastasis, confers an ability of these cells to invade adjacent tissues and to develop into secondary tumors at distant sites.

Little is known about the biochemical or molecular biologic mechanisms governing these important tumor cell traits. It is not even clear that specific activated genes are responsible for mediating metastasis or immune evasion. We suspect that such genes exist, and that they may properly be included under the rubric of oncogenes, because they probably make important contributions to tumorigenesis. A major problem concerns the nature of these genes and how they are able to produce the variety of traits associated with cancer cells.

PARTICIPATION OF ONCOGENES IN THE MULTIPLE STEPS OF CARCINOGENESIS

Carcinogenesis is believed to involve a series of independent, sequentially occurring steps that enable a cell and its descendants to evolve toward malignant growth. Certain carcinogens, termed "initiators," trigger the process of tumorigenesis. Others, called "promoters," facilitate the development of "initiated" cells into full-blown tumor cells. The effects of initiators and promoters on oncogenes are poorly understood. Several of the early steps may involve activation of various growth-promoting oncogenes. Other alterations may propel the cell toward invasion of surrounding tissue and spread to distant sites.

In one very specific case, that of rat embryo fibroblasts, at least two independently activated oncogenes are required to convert normal cells into tumor cells. Although the generality of this observation remains to be demonstrated, the existing evidence underscores the potential complexity, at the molecular level, of the tumorigenic process. The nature and number of molecular events that transpire in the creation of most human tumors are unknown.

How Many Potential Oncogenes Are Harbored by the Cellular Genome?

The cellular genome carries 50,000 or more distinct genes. Of these, only a small portion are concerned with growth control and with other traits expressed by cancer cells. The number of known proto-oncogenes found in the cellular genome now exceeds 20. These have been discovered by virtue of their association with various retroviruses, or by transfection, or by use of cloned DNA probes derived from the first two approaches. It is likely that additional oncogenes remain to be discovered, perhaps by use of new detection procedures.

How Are Oncogenes Activated From Normal Proto-Oncogenes?

An important question concerns the multiplicity of molecular mechanisms that can mediate the activation of proto-oncogenes into their oncogenic versions. For example, as already discussed, the *ras* oncogenes can be activated by discrete point mutations affecting single nucleotides. Amplified versions of *ras* genes are also found in some human malignancies. Another activated oncogene, the *erb*-B gene, results from the truncation of the gene that encodes epidermal growth factor receptor. As a third example, the *myc* gene becomes activated following chromosomal translocation. The precise mechanisms or agents (such as chemical mutagens) that may lead to these changes must be explored. Although much progress has been made in studying environmental carcinogens and their mode of activation within cells, it is still not clear precisely how many of these activated compounds are able to create oncogenes or induce human cancers.

What Might Reverse or Inhibit Oncogenes?

The panel believes that oncogenes lie at the heart of the cancer process and that they represent the centrally acting agents of cellular transformation, rather than peripheral secondary or tertiary consequences of this process. Much oncogene research has been directed toward understanding the causative mechanisms of cancer, but the discovery of these genes also provokes interest in eventual development of anticancer therapeutics. Reversal of the action of oncogenes or their encoded proteins may allow a reconversion of tumor cells to normal cells. Because certain oncogene proteins are structurally different from their normal counterparts, agents might be developed that specifically recognize and antagonize these proteins.

A major challenge is the development of pharmacologic antagonists of oncogene action. Such antagonists would provide an alternative to the cytotoxic agents currently used in cancer therapy. An alternative therapy might depend upon immunologic reagents (e.g., antibodies), which would recognize and destroy cells carrying oncogene proteins. Unfortunately, such an approach may be extremely difficult because many oncogene proteins are located deep within the cell and thus are hidden from immune mechanisms that are able to survey only the cell surface.

An important, relatively unexplored, possibility concerns the natural mechanisms that the organism employs to check cellular growth. Just as there are growth-stimulating factors, there may be corresponding growth-retarding factors. The expression of proto-oncogenes may normally be counterbalanced by growth-antagonizing "anti-oncogenes," which would be valuable in cancer prevention or treatment.

What Types of Genes Confer Hereditary Susceptibility to Cancer, and How Are These Genes Related to Oncogenes?

Many people have inborn susceptibilities to certain types of cancer. In some cases, the genetic factors conferring susceptibility exert their effects by depriving the individual of normal immune defenses for protecting DNA from accumulating large numbers of mutations. But in most cases, the susceptibility is less general, predisposing to specific types of cancer.

The bases for these susceptibilities are not known, but their genetic components are different from the oncogenes described here. The susceptibility genes are associated with an individual from the time of fertilization of the egg; oncogenes are not genetically transmitted in egg or sperm but develop instead in specific tissues during an individual's lifetime. Several of the susceptibility genes have been associated with specific chromosomal aberrations. The identity of these genes and the nature of their functions are obscure at present.

POTENTIAL AREAS OF RAPID DISCOVERY

GROWTH FACTORS AND SECOND MESSENGERS

The growth of most normal cells in the body seems dependent on growth factors, hormones that are transmitted from cell to cell and provide the external stimuli that trigger growth. Cancer cells often can grow in the absence of the normally required exogenous growth factors. From this perspective, oncogenes can be viewed as agents that confer on cells a growth pattern that is independent of external growth factors.

For example, one oncogene has been found to cause the cell itself to make inappropriately large amounts of a growth factor so that the cell's growth becomes independent of exogenous factors. Another oncogene encodes the structure of an altered growth factor receptor in the cell; this altered receptor may make the cell behave as if growth factors are present even when they are not.

These data and hypotheses have resulted in a merging of two previously separate fields of cancer cell biology, oncogenes and growth factors. An understanding of the mechanism of action of growth factors and their receptors will be central to understanding the physiology of the cancer cell.

Ion Channels and Second Messengers

Regulation of cell growth has often been studied by depriving normal cells of growth stimulatory factors and then inducing them to grow by exposure to these factors. Within seconds of exposure to growth stimulatory factors, a variety of changes are seen in the cell. To cite some examples, Na^+ and H^+ are exchanged across the outer membrane of the cell, the intracellular pH rises, free calcium in the cell rises dramatically, and several of the phospholipids of the membrane are degraded to new compounds.

These phenomena illustrate the activity of so-called "second messengers" — compounds in the cell whose changing concentrations transmit important signals from one cellular constituent to many others. It is highly likely that oncogene proteins and their normal versions modulate the concentrations of these second messengers, which then convey growth regulatory information to various parts of the cell. (It may be that some of the ion-regulating mechanisms that originally evolved for mediating cellular growth control were subsequently adapted for modulating ionic states of nerve cells during conductance of excitatory signals.)

These various considerations indicate that two areas of rapid development over the next several years involve study of intracellular second messengers and study of proteins that regulate movement of ions across membranes.

BIOCHEMICAL AND STRUCTURAL ANALYSIS OF ONCOGENE PROTEINS

Study of oncogene proteins has been hampered because only small amounts of material were available. Successful techniques for cloning the oncogenes and expressing their proteins in bacteria have made increased amounts of these proteins available for study. It soon will be possible to conduct structural analyses of these proteins by means of directed modification of the proteins themselves and x-ray crystallographic analysis. Moreover, the recently developed techniques of directed DNA mutagenesis (whereby specific nucleotides can be altered) will make it possible to alter oncogene sequences and thereby redesign specific portions of the proteins they encode.

Use of synthetic oligopeptide immunogens and monoclonal antibodies will enormously facilitate the localization of these proteins to specific cellular sites, the identification of other proteins with which these proteins interact, and the characterization of complex, multicomponent reactions in which these proteins participate.

GENETIC ANALYSIS IN YEAST AND FRUIT FLIES

Genetic analysis in well-studied lower organisms provides a powerful means of looking at properties of proto-oncogenes and oncogenes and their role in cellular and organismic physiology.

In yeast, *Saccharomyces*, and to a lesser extent in the fruit fly, *Drosophila melanogaster*, it is possible to inactivate a proto-oncogene genetically and to study subsequent effects on the organism's viablity, growth, and development. Identification of other genes that interact with oncogenes will provide a new dimension to the current understanding of cellular pathways that are perturbed by oncogene action. Such genetic manipulation is impossible with mammalian cells. Implicit in these experiments is the notion that the molecular mechanisms governing cellular replication in yeast and flies are similar to those in mammalian cells.

The oncogenes discussed until now encode protein products that induce the cell to proliferate. The action of these oncogenes may be balanced by other genes that function to inhibit growth. If such antigrowth genes exist, inactivation of these genes would physiologically mimic the action of oncogenes that actively promote growth.

Some data suggest that certain genes that confer a hereditary, familial predisposition to cancer may act by failing to inhibit growth. Thus, individuals born with an inadequate complement of these "antigrowth" genes may have a lifelong risk of increased cancer incidence. How can we look for these "antigrowth" genes or, more precisely, for the absence of them? It should be possible to use genetics to identify growth-restraining genes in yeast and to employ these yeast genes to find counterparts in mammalian cells.

A further use of genetics derives from our newly gained ability to introduce genes into the germ lines of *Drosophila* and mice. Studies in which oncogenes are introduced into these germ lines will allow researchers to look at the effect that these genes have on normal development of the embryo or on specific organs within the embryo. Such studies promise to provide important new insights into the roles of (proto-) oncogenes in the development of complex organisms and the interaction of certain oncogenes with specific tissues.

The Search for New Oncogenes

Discovery of additional oncogenes is limited by the techniques currently used to detect them. New assays for oncogene activity will be developed over the next several years and presumably will reveal oncogenes whose presence in human DNA was not apparent using existing techniques.

Certain types of oncogenes should be discovered by virtue of their ability to transform indicator cells other than the currently used NIH3T3 mouse fibroblasts. Other oncogenes may be found following direct manipulation of the normal human DNA. Yet others may be found by virtue of their association with specific chromosomal rearrangements.

Cancer Diagnosis

Several reagents soon will be available to detect oncogenes and distinguish their products from those of normal counterpart genes. These reagents may allow earlier identification and greater specificity in detecting tumors than is currently possible.

The success of such diagnosis is predicated upon an issue as yet unresolved: How often are altered oncogenes involved in the causation of human tumors? About 20 percent of human tumor DNAs yield active oncogenes upon transfection into NIH3T3 cells. However, this assay has limitations. In fact, we already know that some oncogenes such as *myc*, which was identified because of its involvement in chromosomal translocations, do not affect NIH3T3 cells and would have been overlooked in such an assay. In other cases, activation of a proto-oncogene may be a necessary but not sufficient step in carcinogenesis and may not warrant a diagnosis of an actively growing cancer.

As methods are improved for discovering new oncogenes, it is the panel's opinion that all human tumors eventually will be found to carry cellular oncogenes. Because oncogenes are important determinants of cellular phenotype, we expect that oncogene diagnosis will provide important prognostic indicators for the clinician dealing with specific tumors. Oncogene diagnosis also may permit better delineation among cancer types and provide better clues to external triggering agents.

NEED FOR CONTINUED BASIC RESEARCH

The great amount of progress in oncogene studies has been possible because of the basic biologic research that preceded and led to the discovery of oncogenes. Continued study in fields such as membrane biology, signal transduction, cellular differentiation, molecular genetics, and cellular immunology will come to bear on the problem of cellular transformation and tumorigenesis. Conversely, the lessons learned from oncogene research will have important implications for many of the unsolved problems of modern biology.

Although the study of oncogenes cannot be expected to resolve all the diverse phenomena that constitute the biology of cancer, we believe that the progress to date has already placed us close to a clear understanding of many of its central mechanisms. Further progress will depend on continuing support of the basic biologic disciplines whose successes have proved so useful in confronting a multitude of biologic problems, cancer being only one of them.

SUMMARY

Within the past 5 years, molecular and cell biologists have made great advances in identifying the molecular mechanisms of carcinogenesis. They have found specific genes within tumor cells, termed "oncogenes," which are responsible for creating many of the abnormal traits of cancer cells. These discoveries provide the foundation for a complete understanding of the molecular basis of cancer within the coming decades.

While these genes have been assigned important roles in carcinogenesis, their precise mechanism of action remains unclear. Much work will be pursued on the structure and enzymatic functions of the oncogene proteins. Thus, certain oncogenes may affect the levels of important compounds in the cell that regulate growth. Others may regulate the expression of certain cellular genes. Yet others may regulate the expression of secreted growth-regulating hormones or receptors used by the cell to detect these hormones.

Future advances in understanding oncogene function will depend on exploiting a variety of experimental tools made available by such diverse basic biologic disciplines as pharmacology, molecular biology, genetics, and cell biology. The presence of related genes in organisms such as fruit flies and yeast will make it possible to study the oncogenes in the context of these genetically simpler organisms. Future advances also will focus on how chemical carcinogens can activate oncogenes and how other agents may antagonize oncogene function and the growth of the cancer cell. New genes may be found that function to control oncogenes and inhibit their functions. Such advances may lead over the next decades to insights into novel therapies that interfere specifically with the growth of cancer cells.

Report of the
Research Briefing Panel on
Interactions Between Blood and Blood Vessels
(Including the Biology of Atherosclerosis)

Research Briefing Panel on Interactions Between Blood and Blood Vessels (Including the Biology of Atherosclerosis)

Alfred P. Fishman (*Chairman*), Director, Cardiovascular-Pulmonary Division, Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pa.

Ramzi S. Cotran, Professor of Pathology, Harvard Medical School, and Pathologist-in-Chief, Brigham and Women's Hospital, Boston, Mass.

Harriet P. Dustan, Director, Cardiovascular Research and Training Center, University of Alabama in Birmingham, Birmingham, Ala.

Howard A. Eder, Professor of Medicine, Albert Einstein College of Medicine, Bronx, N.Y.

Judah Folkman, Professor of Pediatric Surgery and Professor of Anatomy, Harvard Medical School, and Senior Associate in Surgery, Children's Hospital, Boston, Mass.

Robert Furchgott, Professor of Pharmacology, Downstate Medical Center, Brooklyn, N.Y.

John C. Hoak, Professor and Chairman, Department of Medicine, University of Vermont, Burlington, Vt.

Eugene Renkin, Professor and Chairman, Department of Human Physiology, School of Medicine, University of California, Davis, Calif.

John Repine, Professor of Medicine and Pediatrics, Webb-Waring Lung Institute, Denver, Colo.

Russell Ross, Professor and Chairman, Department of Pathology, School of Medicine, University of Washington, Seattle, Wash.

Una Ryan, Professor of Medicine, University of Miami School of Medicine, Miami, Fla.

Alan D. Schreiber, Professor of Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pa.

Thomas W. Smith, Chief, Cardiovascular Division, Brigham and Women's Hospital, and Professor of Medicine, Harvard Medical School, Boston, Mass.

Peter Ward, Interim Dean, University of Michigan Medical School, Ann Arbor, Mich.

Albert Winegrad, Professor of Medicine, Cox Institute, Philadelphia, Pa.

Staff

Barbara Filner, *Director*, Division of Health Sciences Policy, Institute of Medicine

Barbara Mandula, *Consultant*

Naomi Hudson, *Administrative Secretary*

Report of the Research Briefing Panel on Interactions Between Blood and Blood Vessels (Including the Biology of Atherosclerosis)

In atherosclerosis (a form of hardening of the arteries), blood constituents are deposited in the blood vessel wall, causing clotting and inflammation. If the vessel becomes blocked, the consequences can be extremely serious. Medical care and medical research have long been concerned with the complications of atherosclerosis, ranging from heart attacks and strokes to peripheral vascular insufficiency and kidney disease. Complications of atherosclerosis are relatively common, account for an extraordinary toll of chronic impairment and early deaths, and result in staggering economic costs. Despite a 30 percent decline in death rates from heart attacks and strokes in the United States in the past 15 years, complications of atherosclerosis still cause almost half of all deaths in this country. The disease was identified as the leading cause of days lost from work in 1977 and was estimated to have cost almost $40 billion in health expenditures and lost productivity in the United States.

It is now evident that atherogenesis, the disease process that culminates in atherosclerosis, entails a complex interplay between blood and the vessel walls that contain it. In recent years, great progress has been made in understanding this interplay, and this should enhance our ability to treat and prevent the diseases that result from abnormal blood/vessel interactions. The new knowledge has resulted from interdisciplinary efforts drawing on fresh concepts and new techniques of molecular and cell biology, physics, chemistry, and genetics.

Atherogenesis has been recognized as just one example, albeit one of overriding practical importance, of the disease processes that occur when there is a breakdown of normal interactions between blood constituents and the vessel wall. Studies of such interactions also pull together information about normal and abnormal functions of the circulation, the lungs, and the heart. These studies have the potential to provide useful information for a wide variety of human ills.

This paper considers first the characteristics of the blood vessel wall, especially the endothelium, which is in contact with the blood. Second, the paper discusses the properties of some components of the blood that interact with vessel walls. Third, examples are provided of interactions between blood components and the vessel wall that are of broad biomedical significance. Each of these examples is intended to illustrate normal interplays and the consequences of their disordering by disease or injury. Atherogenesis will be discussed in some detail in the context of recent findings. Finally, opportunities that require further investigation are identified.

THE CIRCULATORY SYSTEM

Each component of the circulatory system has its own distinctive character, metabolism, behavior, and disorders. Blood flowing through this system consists of a clear liquid plasma in which distinctive cells are suspended. The large vessels serve as conduits, but the tiny vessels, principally the capillaries, serve as semipermeable barriers for exchange, between blood plasma and tissues, of oxygen, carbon dioxide, nutrients, wastes, and other substances. Capillaries also have functions peculiar to the tissue or organ in which they are situated. In the lungs, for example, the vast capillary network acts as a physical filter to prevent particles, such as clots, from reaching and harming vital organs, particularly the brain and heart.

Large arteries are subject to atherosclerosis and both arteries and veins incur thrombosis, or clotting. Capillaries can be damaged by inflammatory diseases, such as occurs in the kidney in glomerulonephritis.

Although the architecture and basic structures of the circulatory system have been known for a long time, the detailed molecular mechanisms by which the system carries out its functions are only now being unraveled. The use of modern biologic techniques has made it possible to analyze interactions among the various components of the system and to accrue new information about control mechanisms in health and disease.

ENDOTHELIUM

The inner lining of blood vessels is a single layer of cells, the endothelium, which has remarkable properties. It has an apparently simple structure and once was thought to fulfill fairly simple tasks; now it is known to have biologic functions that are many, complex, and sometimes unique to the organs in which the blood vessels are located. In general, the monolayer serves as a selective barrier to the movement of fluids and cells between blood and tissue, as an antithrombotic lining, and as a metabolic organ in its own right.

The lining of vessels normally allows the passage of blood for a lifetime without clotting and without its constituents becoming denatured. No artificial substitute has yet been created that can do that. But an injury upsets the dynamic balance between the vessel wall and blood constituents. A subtle injury, in the form of sustained high blood pressure or high blood cholesterol levels, appears to enhance the permeability of the vascular wall to plasma constituents and enables cells to migrate from the blood stream and possibly contribute to the beginnings of disease.

Information transfer among the components of the vessel wall, and between the endothelial lining and blood constituents, is prerequisite for normal structure and function of the circulatory system. In capillaries, whose wall consists entirely of endothelium, communication between contiguous cells somehow ensures that the vessel lining is a continuous monolayer. In arteries, whose walls are more complex, endothelium communicates with muscle cells that form the next layer outside of it, generally influencing the state of contraction or dilation of the vessel. How this cell-to-cell communication operates is of great investigative interest, and the results of current studies may radically change traditional concepts of how blood flow and pressure are controlled.

Endothelium as a Metabolic Organ

New techniques of culturing endothelial cells have made it possible to study their biochemical activity under controlled conditions and to harvest sufficient material for biochemical study. A combination of in vivo and in vitro studies, for example, has disclosed the vital role of pulmonary capillary endothelium in maintaining systemic blood pressure. On the luminal surface of pulmonary capillary endothelium, the surface in contact with blood, angiotensin-converting enzyme degrades bradykinin (a vasodilator) to inactive products. The same enzyme also converts angiotensin I to angiotensin II (a vasoconstrictor), thereby further helping to sustain systemic blood pressure. In addition, pulmonary capillary endothelium generates angiotensin III, a hormone that acts on the kidneys to influence blood pressure by affecting salt and water balance.

Endothelial cells also participate in maintaining the fluidity of the blood while ensuring that clot formation occurs locally as needed to limit bleeding. Sequences of enzymatic reactions ("cascades") have been identified that either form clots or dissolve clots. Cellular components of the blood, particularly the platelets, also contain powerful substances that lead to clot formation when they encounter injured endothelial lining. The endothelial lining locally generates a hormone (prostacyclin) that modulates the clotting response. Bradykinin, released and destroyed locally, enhances the generation of prostacyclin. These are complicated interplays, regulating the coagulation-anticoagulation systems. Much remains to be learned about the coordinated responses, the mechanisms that prevent under- and overreaction during injury or stress, and the imbalances that arise from kidney failure, or a variety of prescribed drugs, or inherited disorders involving these components.

The growing understanding of endothelial functions has been paralleled by insights into its structural complexities. Ultrastructural specialization of the endothelial cell surface occurs, allowing it to localize diverse functions and carry them out more effectively. Some domains of functional activity are distinctive because of electrostatic charge, while others occur as invaginations in which enzymes are concentrated. How the various molecular activities are organized and orchestrated is currently being explored in depth.

Further study is needed to compare the behavior of endothelial cells in vivo and in vitro, to determine optimal methods for growing endothelial cells from different sites and organs in vitro so as to simulate natural conditions, and to compare endothelial function in large and small vessels. Endothelial cells in culture also are beginning to be used to produce biologically useful substances, such as plasminogen activator, which is important in clot dissolution, and to test pharmacologic agents for toxicity. This research also should continue to be productive.

Transport of Material Across Endothelial Cells

Substances must cross the endothelium in order to leave the blood and be delivered to tissues. Scientists are using a variety of techniques to explore the mechanisms involved in the passage of substances between blood and tissues and the reasons for the different structures of endothelium in different parts of the body.

Endothelium that lines the capillaries of muscle and lungs is continuous; the cells of the monolayer are linked by tight junctions. In contrast, endothelium that lines liver sinusoids and kidney glomerular capillaries has gaps between adjacent endothelial cells so that the adjacent basement membrane, rather than the endothelial layer, constitutes the initial barrier between blood and tissue.

Until recently, the passage of large molecules, such as proteins, was explained in terms of a "pore theory," which envisioned movement through holes or tunnels in the capillary wall. However, experimental observations have made this theory less convincing. At present, there is particular interest in the biochemical attributes of the endothelial cell and the underlying interstitial space as determinants of the transcapillary passage of blood constituents. In addition, attempts continue to relate anatomical structures, such as vesicles, to the transcapillary passage of charged and uncharged particles. The rapidly developing field of transendothelial transport has added fresh impetus and direction to these studies.

Communication Between Endothelium and Smooth Muscle

The state of contraction of the smooth muscle in the walls of small arteries and arterioles largely determines blood pressure and is responsible for adequately distributing blood to all parts of the body. A variety of natural regulatory mechanisms automatically adjust the state of contraction, or vascular tone. Some pharmacologic agents act directly to increase or decrease vascular tone, that is, to cause vasoconstriction or vasodilation.

Only recently has it been demonstrated that certain agents act as vasodilators by acting on the endothelium rather than on the muscle of the vessel wall. It has been proposed that a humoral substance, "endothelium-derived relaxing factor" (EDRF), mediates this dilation. Biochemical substances that seem to exert a vasodilating effect through EDRF include acetylcholine, substance P, adenosine diphosphate, and adenosine triphosphate. The finding has important implications for local regulation of blood supply both under normal conditions and under conditions of vascular disease in which physical and chemical links between cells of the vessel wall are often markedly deranged.

The chemical identity of EDRF has not yet been established. It is possible that EDRF has a central role in maintaining proper blood pressure and that further study of it will help in clarifying and controlling abnormalities in blood pressure.

Angiogenesis

The growth of new capillaries is greatly dependent on the activity of endothelial cells. The endothelial cells that line capillaries are usually in a resting state. Their rate of proliferation is so low that turnover times are measured in years. However, these endothelial cells are capable of sudden growth spurts during periods when new capillaries are rapidly forming, for example, during wound healing. The process of new capillary formation is called "angiogenesis," and cell turnover times during angiogenesis are measured in days.

Protracted angiogenesis is observed in such diverse pathologic states as arthritis, psoriasis, diabetic retinopathy, and chronic inflammation; the most sustained and intense angiogenesis is associated with solid tumors. In many of these abnormalities, angiogenesis itself contributes to the disease process. In arthritis, new capillaries can invade and destroy joint cartilage. In diabetes, new capillaries in the retina of the eye hemorrhage and cause blindness. Progressive growth of tumors and their ability to form metastases depend upon continuous angiogenesis.

Because of techniques introduced since 1970, it is now possible to investigate angiogenesis experimentally. Capillary endothelial cells have been cloned and grown in vitro, where they form tubes, branches, and whole capillary networks. Special polymers have been developed for the sustained release of angiogenic factors in vivo. Bioassays for angiogenic activity have been perfected in the chick embryo and in the rabbit and mouse cornea.

Application of these techniques is revealing some components of the angiogenic process. As an early step in capillary formation, endothelial cells release specific enzymes that punch holes in the vascular basement membrane. Endothelial cells migrate through these holes. The cells align in a linear configuration to form a new sprout. The sprout elongates through interstitial tissue toward an angiogenic stimulus such as a small tumor. Sprout elongation depends upon chemotaxis of endothelial cells at the tip of the sprout, division of cells farther back, and production of interstitial collagenases, probably by all of the endothelial cells in the sprout.

Techniques developed primarily for the study of tumor angiogenesis have recently been used to explore the mechanisms of angiogenesis associated with processes such as wound healing. Endothelial cells and debris-scavenging macrophages play major roles. In the depth of a fresh wound, tissue oxygen levels are extremely low because capillaries have been damaged and blood flow interrupted. In a setting that is poor in oxygen, macrophages that migrate into the wound secrete angiogenesis factors. As new capillaries are attracted into the wound, oxygen tension rises; the macrophages then reduce or shut off their stimulation of angiogenesis. The details of this feedback mechanism remain to be worked out, and the macrophage angiogenic factor has not yet been purified.

Blood Constituents

Blood is composed of circulating cells suspended in a liquid, the plasma. Each component of blood has access to endothelium and virtually all bear electrostatic charges that interact with each other and with the surface charges on the endothelial lining. The composition of the blood is automatically adjusted to the various needs of the body, including respiratory gas exchange, maintenance of blood pressure, and maintenance of fluid and electrolyte balance. Blood constituents interact with each other as well as with the vessel wall. A few of the blood constituents that have crucial roles in interacting with the vessel wall are discussed below.

Platelets

Platelets are essential elements in blood clotting and also have a central role in the process of atherogenesis. Upon contact with collagen or macrophages in an area of vascular injury, platelets aggregate and their storage granules discharge biologically active substances that promote the formation and organization of a clot to plug the breach in the endothelial lining. Enzymes and growth factors released by the aggregated platelets attract other platelets, evoke spasm and proliferation of adjacent vascular smooth muscle, and draw coagulation proteins to the area for the assembly of the clot. These reactions are controlled by molecular feedback mechanisms. For instance, the activity of thromboxane, released by platelets to promote local vascular constriction and accumulation of other platelets, is modulated in its clot-promoting effects by prostacyclin (an anti-aggregating and vasodilator agent) liberated by endothelium.

New techniques have greatly expanded our knowledge of platelet structure and function. The surface of the platelet displays specialized molecular organization, which includes receptors that provide binding sites for macromolecules that engage in complex immunologic reactions, and other receptors that have to do with the clotting process and that bind fibrinogen and von Willebrand factor. These receptors constitute ultrastructural foci for the orderly assembly of the clot. As the coagulation proteins fall into place on the platelet surface, the clotting process is dramatically accelerated by a surge in formation of thrombin.

There are still large gaps in the understanding of how platelets engage in their complex interactions with the vessel wall, plasma factors, and other blood cells. Their role in the inflammatory process, for example, remains enigmatic. Platelets, in addition, are a storehouse for powerful enzymes that differ strikingly in structure and function. How they interrelate in response to injury requires much further study.

Leukocytes

Leukocytes, or white blood cells, are part of the body's defense system. They are drawn chemotactically to areas of damage or invasion by foreign agents and participate in the ingestion, walling off, and elimination of the offending agent. Upon stimulation by appropriate chemicals, leukocytes generate oxygen radicals and release substances capable of destroying microorganisms. We are now beginning to understand how the body protects itself against inappropriate activity of these powerful chemicals. Phagocytic cells involved in the inflammatory response, as well as other cells of the body, are equipped with intracellular antioxidant enzymes, such as catalase and glutathione peroxidase; in addition, there are circulating neutralizers, such as ceruloplasmin and alpha-1 antitrypsin, to prevent these chemicals from damaging the host when they are released. It is anticipated that awareness of how leukocytes inflict their damage will make it possible to develop medications to protect people against their toxic products.

Plasma Lipoproteins

Because lipids are insoluble in an aqueous medium, they are transported in plasma as macromolecular complexes called lipoproteins. Of the various types of lipoproteins, low-density lipoprotein (LDL) carries one-half to two-thirds of all the cholesterol in plasma. LDL has cholesterol ester as its core lipid and unesterified cholesterol on its surface. High levels of LDL-cholesterol predispose to atherosclerosis. But another form of lipoprotein, high-density lipoprotein (HDL), may be anti-atherogenic.

Cells of many different tissues have specific receptors for LDL and transport it into the cytoplasm, where it subsequently is degraded in lysosomes. In normal circumstances, however, the entry of LDL into the artery wall is predominantly receptor-independent and occurs by a nonspecific process known as transcytosis. Reverse transport of LDL also occurs and avoids undue accumulation of cholesterol within the cell. The mechanisms involved in reverse transport and in the feedback control of intracellular cholesterol concentrations are under investigation, as is the mechanism of LDL transport into and out of endothelial cells following injury to the vascular lining.

Macrophages and endothelial cells have specific receptors for LDL, although their rate of uptake of LDL is low. Macrophages also have specific receptors for LDL that has been chemically modified or altered by contact with endothelial cells. The rate of uptake of modified LDL by this receptor is considerably higher than that of unmodified LDL, and foam cells (macrophages filled with cholesterol) can be readily produced by interaction of modified LDL with macrophages. LDL also apparently alters endothelial cell membranes in a way that allows access of lipoproteins into the vessel wall. When LDL is present in the vessel wall, it may cause proliferation of smooth muscle cells, which migrate into the lining of the blood vessel of the affected artery.

Considerably less is known about the interaction of vessels with the larger lipid-rich lipoproteins, such as chylomicrons. These are triglyceride-rich particles that appear in the circulation after the ingestion of fat and in certain disease states. These are removed by the liver, which then secretes the cholesterol in the particles in the form of VLDL, which is metabolized in part to LDL. It is by this mechanism that dietary cholesterol eventually enters the vessel wall.

In persons with a disorder called Type III hyperlipoproteinemia, atherosclerosis occurs frequently. In these patients, the chylomicron remnants and VLDL remnants enriched in cholesterol accumulate in the plasma in a form called beta-VLDL. Macrophages from patients with Type III hyperlipoproteinemia have receptors for beta-VLDL, which probably accounts for the accumulation of foam cells in vessel walls in the distinctive skin lesions of these patients.

Earlier studies had suggested that high-density lipoproteins (HDL) are involved in transporting cholesterol out of macrophages. HDL contains a number of apoproteins (the protein portion) of lipoproteins. The major one in HDL is apo A-I. A specific receptor for apo

A-I has been found on endothelial cells, on vascular smooth muscle cells, and on some liver cells. It has been demonstrated recently that the apo A-I receptor on nonliver cells is regulated by the cholesterol content of the cells. Thus, HDL may transport intracellular cholesterol out of the vessel, becoming enriched in cholesterol while doing so. How the receptor for HDL on liver cells is regulated is not yet known.

Mechanisms of entry of LDL into the vessel wall and handling of LDL when endothelial cells are damaged require further investigation. The extent of interaction of chylomicron remnants or VLDL remnants with the vessel wall also needs study. An exciting new area of investigation involves the mechanism by which HDL, which can bind to specific receptors on vascular endothelial cells and on smooth muscle cells, transports cholesterol from these tissues. New techniques, such as production of monoclonal antibodies for the various apolipoproteins, will be important for studying the metabolism of apolipoproteins and their role in atherogenesis.

EXAMPLES OF BIOMEDICALLY IMPORTANT INTERACTIONS

Molecular interactions between the constituents of blood and the endothelium operate continuously to permit normal growth and development and to sustain the metabolic activities of daily life. Respiratory gases, water, nutrients, and other molecules must move between endothelium, plasma, and the cells of the blood. Electrolytes pass back and forth between blood and tissues where they trigger and drive metabolic machinery. Hormones travel long distances in blood to activate and participate in biological processes far removed from their sites of origin. Each of these systems is governed by elaborate control mechanisms that continuously monitor and restore dynamic equilibria. To illustrate the operation, complexity, and control of the biologic systems and their regulatory mechanisms, three examples will be considered: blood clotting, the inflammatory process, and atherogenesis.

Blood Clotting (Thrombosis)

The tendency for blood to clot when a blood vessel is severed is a safeguard against bleeding to death. In living systems, this tendency is counterbalanced by mechanisms to ensure blood fluidity; inappropriate or excess clotting may lead to serious clinical syndromes. Thrombosis plays a critical role in initiating, propagating, and occluding the atherosclerotic lesion, thereby constituting the precipitating cause of acute myocardial infarction, stroke, and peripheral vascular disease. Thrombosis is a precursor of thromboembolic disease, a common and life-threatening complication of surgery, heart failure, and cardiac arrhythmias. Improved understanding of thrombotic events—involving endothelium and blood constituents—holds the prospect of better therapeutic and prophylactic interventions.

The complex sets of reactions involved in clot formation and removal have been examined experimentally for years, but more pieces of the puzzle continue to be found. For example, one recent discovery is Protein C, a vitamin-K-dependent coagulation factor and a potent anticoagulant and fibrinolytic agent. Its concentration is regulated by an elaborate biochemical control system: Protein C is activated by thrombin bound to the protein thrombomodulin on endothelium; activated Protein C is neutralized by an inhibitor present in the plasma. Protein C appears to be crucial for life and several infants who lacked Protein C because of a genetic defect died of generalized thrombosis shortly after birth.

The Inflammatory Process

The inflammatory process is a vital component of the defense mechanisms of the body. A cellular response, abetted by an increase in endothelial permeability evoked by inflammatory peptides and lipids, contributes to the leakage of water, salts, and proteins from plasma into the extravascular environment. Inherent in this process is the migration of leukocytes from the blood to the site of injury. En route, these blood cells traverse endothelium.

New concepts and techniques are currently being applied to the interplay between the blood and endothelium in the inflammatory process. Promising lines of research include the effects of particular types of endothelial injury on the inflammatory response, the response of endothelium to substances, such as bradykinin, that are involved in the inflammatory response, and the mechanisms that promote adherence between neutrophils (a type of white blood cell) and the endothelial cell.

ATHEROGENESIS

Atherosclerosis is a special type of arteriosclerosis that affects large arteries, such as those of the coronary and cerebral circulations. Atherogenesis is the process that culminates in atherosclerosis. It results from misdirections and overresponses in normal biologic interplay at the molecular level. In its early stages—those most relevant to the understanding of initiating mechanisms and the institution of preventive measures—the inner layer of an affected artery is thickened by a deposition of fat, particularly cholesterol, and by an increase in the number of cells in the lining of the vessel. Two hypotheses concerning the etiology and pathogenesis of atherosclerosis continue to be studied intensively: the lipid infiltration hypothesis and the response-to-injury hypothesis. Each of these has stimulated new knowledge about the interplay between the vascular wall and the blood in atherogenesis.

Lipoproteins in Atherosclerosis

As noted above, the importance of lipoproteins as the cholesterol-carrying substances in blood has been clarified during the last decade. The recent decrease in mortality from heart attacks and strokes has been attributed, in part, to improved understanding of the lipoproteins and their relation to risk factors, such as high blood pressure, smoking, diabetes, and high levels of blood cholesterol. However, still to be resolved is how risk factors act as injurious agents and in what ways the responses to these diverse types of injury differ. New approaches to slowing, arresting, and reversing the atherosclerotic process are being sought. Among the more promising is modification of diet, coupled as needed with the use of medications designed to lower blood levels of cholesterol. One recent lead is the possible effect of certain fats present in seafood in lowering blood levels of cholesterol and low-density lipoproteins.

Response-to-Injury Hypothesis

This hypothesis for atherogenesis, proposed in 1973, suggested that the various risk factors associated with atherosclerosis somehow lead to alterations in the endothelial cells, and that these represent various forms of "injury."

An important development associated with the response-to-injury hypothesis has been the discovery of growth factors that may be important in the development of the lesions of atherosclerosis. Two of these are the platelet-derived growth factor (PDGF) and the macrophage-derived growth factor. PDGF is the principal substance in whole blood serum that is responsible for proliferation of cells, such as smooth muscle and fibroblasts, in cell culture. PDGF may be associated with initiation and progression of the lesions of atherosclerosis, because experimentally induced lesions can be prevented by agents that substantially decrease circulating platelets or inhibit platelet interactions.

PDGF plays an important role both in mobilization and proliferation of cells. It has been purified to homogeneity, characterized, and sequenced. PDGF binds to a specific high-affinity receptor on the surface of susceptible cells, and promptly induces such intracellular events as tyrosine phosphorylation of several proteins, including the receptor for PDGF; phospholipase activation leading to diglyceride formation; and formation of arachidonic acid and prostaglandin. It also leads to increased binding of lipoproteins, increased endocytosis, increased protein and RNA synthesis, and then, after 24 to 36 hours, to DNA replication and cell division. Platelet-derived growth factor also is a chemotactic agent that can attract cells, such as smooth muscle, into the cavity of the affected artery.

Experiments concerning the cellular and molecular biology of endothelium (which is also a potential source of platelet-derived growth factor), smooth muscle, the monocyte/macrophage, and the platelet will be important in understanding the lesions of atherosclerosis. Studies that examine the interactions of these cells have opened new frontiers and represent vitally important areas for further research.

Thrombosis in Atherosclerosis

Advances in our knowledge of blood clotting factors, platelets, and prostaglandins have stimulated new insights into old theories about how thrombosis is involved in all stages of atherosclerosis. Studies have demonstrated that platelets, the principal cells involved in thrombosis, may play a role in the initiation of atherosclerosis. Such studies may help to control and prevent thrombosis.

IMPLICATIONS AND PROSPECTS

The concepts and techniques of cell biology that deal with structure, function, and control mechanisms at the molecular level are now available for application to the important biomedical problem of interactions between the vessel and the blood that flows through it. To date, the major focus has been on the individual components involved in this interplay. The time is now ripe for a concerted interdisciplinary effort to understand the control mechanisms that regulate the interplay in health and the derangements arising from perturbations in the system arising from injury and disease.

Favoring a concerted effort at this time is the current capability for harvesting large quantities of endothelium grown in culture. In addition, innovative cell sorters and isolation techniques make it possible to isolate manageable quantities of individual blood cells for manipulation and study. Biochemical and biophysical techniques also provide precise tools for quantifying constitutents in plasma. Finally, receptors and channels at the level of the cell and its organelles can be examined using approaches of molecular and cell biology that were previously unavailable.

The practical rewards of this fundamental research have been suggested in the sections of this paper devoted to thrombosis, inflammation, and atherosclerosis. Many other prospects can be reasonably expected to follow greater understanding of the blood and blood vessel interactions: improved prosthetic devices, such as heart valves, vascular replacements, dialyzing membranes, and artificial organs; fresh insights into the prevention and treatment of vascular occlusions and their complications; new approaches to the prevention, diagnosis, management, and reversal of atherosclerotic lesions. Study of interactions between the blood and the blood vessels provides a dramatic opportunity for progress in fundamental science and holds extraordinary promise for practical application.

Report of the Research Briefing Panel on the Biology of Parasitism

Research Briefing Panel on the Biology of Parasitism

John R. David (*Co-chairman*), John LaPorte Given Professor and Chairman, Department of Tropical Public Health, Harvard School of Public Health, and Professor of Medicine, Harvard Medical School, Boston, Mass.
Joshua Lederberg (*Co-chairman*), President, The Rockefeller University, New York, N.Y.
George Cross, Professor of Molecular Parasitology, The Rockefeller University, New York, N.Y.
Paul Englund, Professor of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, Md.
Paul Gross, President and Director, Marine Biological Laboratory, Woods Hole, Mass.
Seymour J. Klebanoff, Professor of Medicine, and Head, Division of Allergy and Infectious Diseases, Department of Medicine, University of Washington, Seattle, Wash.
Adel Mahmoud, Professor of Medicine, Chief, Division of Geographic Medicine, Case Western Reserve University, University Hospital, Cleveland, Ohio
Adolfo Martinez-Palomo, Head, Section of Experimental Pathology, Centro de Investigacion y de Estudios Avanzados del Instituto Politechnico National, Mexico City, Mexico

Staff

Barbara Filner, *Director*, Division of Health Sciences Policy, Institute of Medicine
Barbara Mandula, *Consultant*
Naomi Hudson, *Administrative Secretary*

Report of the Research Briefing Panel on the Biology of Parasitism

Parasites are organisms that require another organism, a "host," for survival and that usually do some harm to the host. The field of parasitology has been limited to those organisms that belong to the animal kingdom, although bacteria, viruses, and fungi are also parasites. Of concern here are the single-celled protozoans, such as cause malaria and leishmania; multicellular worms or helminths, such as schistosomes and filaria; and arthropods, which include insects. Many of the arthropods also are vectors that transmit the protozoan and helminthic parasites.

There are two major reasons for the current intense interest in the study of parasites. First, such studies will have a large impact on our understanding of basic biologic processes such as those involved in cell growth and differentiation. Many of the human parasites have multiple hosts and may have several developmental stages in each host. Each stage involves regulated transitions from one form to another. Further, parasites have evolved very specialized adaptations that enable them to infect the host and then evade the host's defense mechanisms while they survive and replicate, utilizing for these purposes the host's metabolic processes. The mechanisms by which parasites adapt to their environment are diverse and ingenious, and are only beginning to be understood.

Second, application of the knowledge obtained in these basic investigations promises to have a major impact on world health. Of the six diseases singled out for research emphasis by the Special Program for Research and Training in Tropical Diseases, sponsored by the World Bank, the United Nations Development Program, and the World Health Organization, five are caused by parasites: malaria, schistosomiasis, filariasis, trypanosomiasis, and leishmaniasis. These and other parasites affect more than a billion people worldwide. Approximately 300 million suffer from malaria, 200 million have schistosomiasis, and 300 million have filariasis. Another 40 million have onchocerciasis, which causes "river blindness."

Parasite diseases also affect the United States. Giardia, a protozoan parasite, is one of the most common causes of epidemic infectious diarrhea in this country. In addition, we must consider the immigration of a large number of people with parasitic diseases, the vast outflow of American tourists and business representatives overseas, and the potential exposure of military personnel in countries where parasitic infections are common.

The potential rewards of studying parasites are enormous, both in terms of learning about basic biologic processes and in combating disease. Because of their complexity, animal parasites are difficult to study. However, modern biologic concepts and advances in techniques are now enabling researchers to address such exciting and long-standing questions about parasite-host interactions as the following: (1) How are these unicellular and multicellular organisms capable of establishing infections? (2) What genes regulate the transformation of parasites through their complex life cycle? (3) What are the mechanisms of disease syndromes that result from parasitic infections? (4) What is the array of possible host protective responses and can these be enhanced? (5) What mechanisms do parasites use to evade the host's immune and other protective defenses? (6) Are vaccines feasible? (7) What biochemical pathways do parasites have that differ from the host and can these be used as targets for new drugs? (8) What is the basis for the ability of certain insects to transmit parasites or to become resistant to insecticides, and can these properties be altered?

Some of the technical advances that have been crucial to the study of parasites are the production of monoclonal antibodies, the isolation of specific genes, and the culturing and growth of some of these organisms in the laboratory. This paper describes some basic biologic questions that are being studied in parasites by several different scientific disciplines, and then suggests potential future applications of such information to prevention and treatment of parasitic diseases.

BASIC RESEARCH ON PARASITES

Molecular Biology

Antigenic Variation

African trypanosomiasis is characterized by cyclic parasitemia, that is, by the appearance in the blood stream every 7 to 10 days of waves of organisms. The surface of the protozoan's membrane is covered by a single type of glycoprotein, which is antigenic, and the parasites of each cycle contain a completely different glycoprotein. Although the body mounts an antibody response, the parasite can change its surface antigen hundreds of times, always keeping ahead of the host's immune response.

What is the basis for the capacity of these parasites to alter their surface antigens? Investigation of this phenomenon, undoubtedly the most sophisticated mechanism yet devised for evading the immune response of the host, uncovered the first example of "jumping genes" for surface proteins. Each organism has an estimated 300 to 1,000 different genes coding for these variant surface glycoproteins, but only one is expressed at a time. The gene to be expressed is duplicated and then moved to another part of the chromosome where it is expressed. It is not yet known how these parasite genes are selected, how they are turned on and off, or how they are regulated. Nevertheless, studies of the molecular biology of the African trypanosome already have had considerable influence on our understanding of gene rearrangement in general, and promise to uncover mechanisms of gene regulation.

Surface Antigen Repeating Units

Sporozoites, the form of the malaria parasite that matures in the salivary gland of the anopheles mosquito and is injected into the human skin during a bite, are free in the human host's blood for only a few minutes before entering liver cells. Because of their precise specificity for a single antigen, monoclonal antibodies have allowed the detection and characterization of surface proteins on sporozoites. The antibodies are specific to a given species as well as to that stage in the life cycle. The antibodies directed against these surface antigens neutralize sporozoite infectivity, so there has been great interest in both the antibodies and the antigens they have identified.

Genes coding for the sporozoite surface antigens have been introduced into bacteria by recombinant DNA techniques. This allows the production of sufficient antigen to study it in detail. The antigens have been found to have unusual repeating sequences. In the sporozoite of the malaria species *Plasmodium knowlesi*, for example, the unit is made up of a peptide with 12 amino acids repeated 12 times in tandem. The antigen of *Plasmodium falciparum* has a different repeating unit, with 4 amino acids repeated 23 times. It will be interesting to learn how common such repeating antigen units are in other parasites and what evolutionary pressures may have led to their development. In a more practical vein, such small reactive units lend themselves to the production of synthetic peptide antigens that may be used to develop vaccines effective against malaria.

Kinetoplast DNA

Kinetoplast DNA (kDNA) is a form of mitochondrial DNA found in some protozoans such as trypanosomes and leishmania, which is essential for their survival. It has a remarkable structure consisting of thousands of DNA circles interlocked in a network. A minor portion of these circles, the maxicircles, code for mitochondrial proteins. The majority, the minicircles, are only about 1,000 base pairs in size. These do not appear to code for any protein, and their function is not known. Minicircles are the only DNA in nature known to have a region of bent DNA helix. It is of great interest to learn both the function of this unusual DNA conformation and how its nucleotide sequence induces the curvature.

Gene Regulation During Life Cycle Transformations

Parasites undergo profound changes during the various stages of their life cycles, and work on the molecular biologic basis of these changes is in its infancy. The next few years should see dramatic progress in work on the molecular basis for developmental transformation in many different protozoan and helminthic parasites.

Leishmania, for example, go from a flagellated protozoan form in the insect host to a smaller form without flagella when in a mammalian macrophage. This change is reversed after the protozoa are taken up by a fly when it bites an infected animal. The amount of tubulin, a structural protein of the cytoskeleton, decreases as the parasite goes from the flagellated to the nonflagellated form and then increases when the parasite returns to the flagellated form. In this case of transformation, the control of the genes for tubulin occurs at the level of messenger RNA processing.

IMMUNOLOGY

Studies in immunology have concentrated in several areas. These include mechanisms the host can muster that the parasite cannot escape, mechanisms of immune evasion, and the role of the immune response in causing tissue damage. The information obtained has been useful in developing better diagnostic reagents and forms the basis for the development of protective vaccines.

Protective Mechanisms of the Host

Studies on host responses to parasites have revealed some novel systems. For example, one such system was found while studying the antibody-dependent ability of host cells to kill schistosomes. These studies led to the first demonstration that the eosinophil, a white blood cell, can act as a killer cell. These findings have now been applied to other areas where it has been shown that eosinophils play an important part in producing tissue damage in certain types of heart disease, in inflammation, and in allergic diseases such as in asthma.

As another example, work on the immune response to ticks has highlighted the role of basophils, another type of white blood cell, and mast cells in producing vascular permeability in the early stages of cell-mediated immune reactions. Furthermore, studies of the schistosome-linked granuloma (an inflammatory nodule) were the first to demonstrate suppression of lymphocyte hormone production by T-cells. (T-cells and lymphocyte hormones are important regulatory components of the immune system.)

Mechanisms of Immune Evasion

Parasites have developed many ways of evading the human host's defenses. Some change their surface antigens, some masquerade as the host by taking up host molecules onto their surfaces, and others simply shed their surface antigens. Some have enzymes on their surface that destroy antibodies. Macrophages are the body's major cell for engulfing foreign particles, and they usually kill the organisms they engulf. Some parasites have devised ways of avoiding the effects of the toxic substances present in the lysosomes of the macrophage, although how they do this is not known. Yet other parasites suppress the host's immune attack.

Although at present most of our knowledge is simply descriptive, future studies should elucidate the mechanisms for accomplishing all these forms of evasion. Information about the effects of parasites on the immune system would undoubtedly enhance our understanding of how the immune system functions in many other types of diseases.

Immunopathology

Studies on the mechanisms of tissue damage in schistosomiasis demonstrated that the host reactions to egg antigens are the main factors producing pathology. Further studies on this process were the first to show that mononuclear cells of the host release factors that promote fibrosis. The mechanisms of pathology induced by many different parasites merit further study. Of particular interest are the ways parasites induce autoimmune reactions such as those postulated in infections caused by *Trypanosoma cruzi*, the mechanism of blindness caused by *Onchocerca volvulus*, the process of sequestration of infected red blood cells leading to malaria damage in the brain, and the underlying basis for destruction of cartilage in mucocutaneous leishmaniasis.

MEMBRANE BIOLOGY AND CELL BIOLOGY

Unusual Features of Parasitic Membranes

The first interactions between the parasite and the host are membrane-membrane interactions. Many parasites have membranes that are different from those in mammalian cells, and considerable effort has been directed at learning about the function and structure of parasite membranes. As a result, several interesting discoveries have emerged.

One example is the finding that the variant surface glycoprotein of African trypanosomes is attached to the membrane by a novel protein-lipid bond. A newly recognized enzyme is involved in destroying this bond. Another unusual finding concerns an enzyme on the surface membrane of *Trypanosoma cruzi*. This organism has a neuraminidase, an enzyme that cleaves off the most important sugar group on the surface of mammalian cells, namely, sialic acid. The role of this enzyme should be studied, especially as it relates to parasite survival and tissue injury. The question of whether similar enzymes are present on other parasites also is of interest.

Parasite-Host Membrane Interactions

Work is being carried out on the mechanism by which malarial merozoites (the mobile infective stage) invade red blood cells. Glycophorin, a glycoprotein on the red blood cell surface, appears to be an important part of the receptor for the parasite. How parasites disrupt the rigid cytoskeleton of the red blood cell so they can enter it is unknown, as is the source for the special membrane that surrounds the parasite once it is in the red blood cell. Furthermore, parasite antigens soon appear on the outer surface of the red blood cell. To reach the surface, parasite antigens must traverse the parasite's own membrane, then the special membrane surrounding the parasite, and finally the red blood cell membrane. The mechanism for accomplishing this voyage is unknown.

Alteration of the red blood cell surface is important in enabling the red blood cells that contain the parasite of *Plasmodium falciparum*, the most common of the four malaria species, to lodge in the vascular bed of certain organs such as the brain. These surface antigens also undergo antigenic variation by unknown mechanisms. The spleen can alter the expression of these parasite antigens on the red blood cell surface, but how it does so is another mystery to be solved.

Of interest is the recent discovery that *Entameoba histolytica*, another protozoan parasite, kills host cells by injecting a protein into the host cell membrane that causes an ion flux and subsequent osmotic lysis of the target cell. These findings have stimulated a search to see whether analogous proteins are involved in cytotoxicity induced by T-lymphocytes.

Clearly, studies on these membrane effects will have ramifications affecting much more than our knowledge of the parasites themselves.

BIOLOGY OF INSECT VECTORS

An attack on the insects that transmit particular pathogens remains a major focus for most programs aiming to control parasitic diseases. Although chemical insecticides provide the major tool for this purpose, their effectiveness has steadily diminished because of insect resistance.

New technologies should lead to revised strategies for combating insect vectors. Recent developments in molecular biology, combined with improved understanding of insect physiology and behavior, should enable us to better understand the mechanisms of insecticide resistance, vector competence—the ability of an insect to transmit a parasite—and the control, feeding behavior, and reproduction of insects. Ultimately, this technology may enable us to genetically modify insect populations in nature, so that the targeted population is less capable of transmitting pathogens or less susceptible to becoming resistant to insecticides.

Such research, both basic and applied, should have far-reaching benefits, affecting not only the transmission of known disease but also important aspects of agriculture.

BIOCHEMISTRY AND PHARMACOLOGY

The very nature of parasitism implies that the parasite has metabolic needs that the host supplies and, therefore, that the parasite must have metabolic pathways that differ from those of the host. Novel metabolic pathways and even new cell organelles are being discovered. These have led to the design of compounds that can inhibit these unique pathways and thus control some parasites and insects without affecting their hosts. A few examples are described below.

Unusual Metabolic Pathways

In general, parasitic protozoans cannot synthesize purines, which are chemical precursors of nucleic acids. They have therefore developed elaborate pathways to utilize purines from the host. Many of these pathways differ from one protozoan to another. Study of these pathways has greatly enhanced our knowledge of purine metabolism in general and may lead to development of drugs that can kill protozoans but are not toxic to mammalian cells. Polyamines also are required for DNA replication and cell differentiation. Trypanosomes and leishmania have simplified metabolic pathways for biosynthesis of polyamines, which may be susceptible to specific inhibition.

Unusual Organelles

Parasites have been found to contain several novel organelles. One of these, the glycosome, found in trypanosomes, contains the enzymes for glycolysis; its discovery encourages research toward development of specific inhibitors. Another novel organelle is the hydrogenosome found in trichomonas. This organelle contains pyruvate kinase and pyruvate dehydrogenase, two enzymes used in producing the ATP needed by the cell. These hydrogenosomes chemically reduce the drug metronidazol to a toxic compound that kills the parasite. The host cells, which lack this organelle, are spared.

The Neuromuscular Junction

The neuromuscular junction of certain helminths contains receptors for GABA (gamma-aminobutyric acid), which also is a neurotransmitter in the human brain. Ivermectin, a drug that enhances the activity of these receptors, has a profound effect on helminths. It causes paralysis that prevents them from moving or eating. It is exceedingly potent, with less than milligrams per acre of pastureland protecting cattle against helminth parasites. Because this drug cannot traverse the blood-brain barrier, it is nontoxic to humans. Further studies using ivermectin have shown that it also acts on the GABA receptor complex of arthropods. Additional studies with this drug should greatly increase our understanding of insect physiology and yield knowledge of GABA receptors in higher organisms.

Juvenile Hormone

Juvenile hormone (ecdysin) is involved in the differentiation and molting of insects and has recently been found in helminths. Further study of the regulation of molting and differentiation of helminths and the role of ecdysin in this process should lead to entirely new approaches to control of some of these parasites.

APPLICATION TO DISEASES

Increased understanding of the basic biology of parasitism should provide many opportunities for combating diseases caused by these organisms. This is especially important now that traditional public health measures that had helped to control some of these diseases are no longer sufficient by themselves. An example of the need for new methods can be seen in the case of malaria. After World War II, there was great hope that this disease could be eradicated by using DDT to kill the mosquitoes in houses and chloroquin to treat infections in people. Indeed, malaria was eradicated from the United States and around the Mediterranean and decreased markedly in Asia and South America. In the early 1960s in Sri Lanka, the number of cases per year fell from more than 1 million to 18. However, within 5 years, there were again almost a million cases, due to a combination of DDT and chloroquin resistance and to the difficulty in maintaining this costly control program. Now malaria has reached serious epidemic proportions in parts of Asia and South America and continues unabated in Africa where it was never controlled due to the behavior of the local mosquitoes, which, instead of landing on the sprayed walls of the house after biting, go outside. Malaria resistant to the present drugs has become an alarming problem.

The belief that parasitic diseases disappear with modernization and industrialization is not always correct. For instance, the prevalance of schistosomiasis has increased hand in hand with the development of hydroelectric dams required for energy and irrigation projects necessary for improved agriculture. The large lakes behind these dams have added thousands of miles of waterfront and have increased the contact between people and the infected snails that transmit this parasite. Another example is the increase in leishmaniasis in the Amazon where people who are expanding towns and building roads come into contact with sandflies that live in the forest and transmit this infection. Applying insecticides to the forests is neither effective nor feasible.

FUTURE RESEARCH OPPORTUNITIES

There is a great need now to broaden the entire base of biologic research on parasites. This is an opportune time because many of the new concepts and biotechnologies make it possible to solve some of the important questions concerning the biology of parasitism. A multidisciplinary approach is required.

Most of the productive research thus far employing the new biotechnologies has been on two parasites, the malaria parasite *P. falciparum* and the African trypanosome of cattle, *T. brucei*. Much less work in molecular biology, for instance, is being done on the African trypanosomes that infect humans, or on other protozoans and the many helminths.

The main reason for the early focus on *P. falciparum* and *T. brucei* is the relative ease, compared with other parasites, with which these two parasites can be cultured and manipulated in the laboratory. There is much excitement over prospects of a vaccine against the infective form of *P. falciparum*, the sporozoite, but little work has been carried out on the other three species of malaria that infect humans. There is also work in progress on the development of vaccines to the other stages of the *P. falciparum* parasite because it is still not clear how effective the vaccine against the sporozoite will be. The sporozoite vaccine has worked in laboratory tests on rodents, but must now be developed for humans. One problem is that the vaccine will have to work completely in the short time that the sporozoite is present in the blood. An effective adjuvant must be identified, one that will allow for a strong enough response for the vaccine to be clinically effective.

It is now important to take advantage of new technologies to study other parasites. This will require establishing cultures of these organisms in vitro, developing suitable animal models, and establishing appropriate life cycles in the laboratory. Such work is time-consuming, may involve many false starts, and requires cooperation of persons in many disciplines and long-term support.

Molecular biology is providing several new methods for generating antigens needed to form the basis of vaccines. These methods include laboratory synthesis of peptides, producing proteins in microorganisms by recombinant DNA techniques, and incorporation of genes coding for protective parasite antigens into the DNA of attenuated viruses used for unrelated vaccines. The essence of the problem is no longer how to produce antigens by recombinant techniques, but which antigens to produce and how to present them effectively to the immune system in clinically acceptable formulations.

Molecular biologic techniques also are being used for better diagnosis. For instance, the nucleotide sequence of the kinetoplast DNA found in certain protozoans varies among the different species. Using DNA hybridization to identify the kDNA, it is now possible to determine in 24 hours instead of 3 months whether a leishmania infection involves a benign or virulent species. Similar techniques should be applicable to other parasites. Furthermore, the production of species-specific monoclonal antibodies against parasites causing malaria, leishmaniasis, schistosomiasis, and other diseases is leading to the development of greatly improved diagnostic reagents and identification of relevant antigens for the development of vaccines.

The study of the genes that regulate the transformation of the parasite through its different stages should result in novel ways of interrupting the life cycle. Studies of the molecular biology of parasites have already broadened our understanding of molecular biology in general. For instance, it was through the study of malaria that an enzyme, mung bean nuclease, was found recently to have the unique property of cutting genes out of DNA without cutting in the coding region, so that genes can be studied and cloned much more easily than before.

Basic immunologic studies of parasites should be expanded to cover several other areas beyond development of vaccines and diagnostic reagents. For instance, the study of the mechanisms of immune evasion should lead to new methods of overcoming the parasites, and those methods should also be adaptable to such other types of organisms as bacteria, viruses, and fungi. Work on immunopathology, especially possible autoimmune mechanisms, is required to assure that the antigens proposed for vaccines are not also those that can induce pathology detrimental to the host. This type of study is important, for instance, in *T. cruzi* infections, because it has been postulated that the heart disease resulting from this infection may be caused by an autoimmune process.

Basic research on the novel biochemical mechanisms that parasites have evolved is currently one of the most neglected aspects of parasitology. Work in this area should lead to the development of new drugs that are needed to treat some of these diseases. For instance, the major drug for treating African trypanosomiasis is Suramin, first synthesized in 1917. It is a very toxic drug and is effective only in the early stages of the disease, before the parasite reaches the brain and produces the symptoms that give the disease its common name, sleeping sickness. There are no drugs to treat *T. cruzi* infection, and those available against the leishmania organisms were first produced decades ago and are quite toxic.

Studies of the biology of insects that carry and transmit parasites should be expanded to include the application of molecular biology. Alternative methods of insect control, including the genetic modification of insect populations in nature, should result. Investigations in this area should not only help control some of these parasite diseases but also should be applicable to problems in agriculture and animal husbandry.

Despite the enormous potential of research on the basic biology of parasites, relatively few scientists are working in this area, and many more should be encouraged to do so. The foregoing examples of present work on the biology of parasitism should make clear the excitement of imminent discovery and the promise that the research not only will reap new scientific knowledge, but also will enable applications of immense practical consequences. Further study offers tremendous opportunity for progress.

Report of the Research Briefing Panel on Solar-Terrestrial Plasma Physics

Research Briefing Panel on Solar-Terrestrial Plasma Physics

Charles Kennel (*Chairman*), University of California at Los Angeles, Los Angeles, Calif.
Jonathan Arons, University of California at Berkeley, Berkeley, Calif.
Roger Blandford, California Institute of Technology, Pasadena, Calif.
Ferdinand Coroniti, University of California at Los Angeles, Los Angeles, Calif.
Martin Israel, Washington University, St. Louis, Mo.
Louis Lanzerotti, AT&T Bell Laboratories, Murray Hill, N.J.
Dennis Papadopoulos, University of Maryland, College Park, Md.
Robert Rosner, Harvard University, Cambridge, Mass.
Frederick Scarf, TRW Systems, Redondo Beach, Calif.

Staff

Donald C. Shapero, *Staff Director*, Board on Physics and Astronomy, National Research Council
Charles K. Reed, *Rapporteur*, Board on Physics and Astronomy, National Research Council
Helene Patterson, *Staff Assistant*, Board on Physics and Astronomy, National Research Council
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

---

Report of the Research Briefing Panel on Solar-Terrestrial Plasma Physics

1 INTRODUCTION

Radio waves were first reflected from the earth's ionosphere—the very edge of space—in the 1920s. Plasma waves were discovered in the laboratory at about the same time. Soon thereafter scientists began to develop a new branch of physics—plasma physics*—to answer the questions stimulated by observations such as those above. Why do laboratory gas discharges exhibit unexpected collective behavior? How do radio waves propagate in the ionosphere? Why does the appearance of sunspots on the sun presage magnetic storms and auroral displays at earth? What roles do magnetic fields play in stars and galaxies? By the late 1940s, physicists had concluded that plasma is a distinct, fourth state of matter, with properties that are unique because cooperative long-range interactions among the charged particles are far more important than collisions. Moreover, it had become clear that, except for occasional cool regions like planets and their atmospheres, most of the universe is in the plasma state.

Modern plasma physics began in the 1950s. The initiation of research aimed at harnessing the energy source of the stars—thermonuclear fusion—and the first launch of an artificial earth satellite catalyzed the development of the technology needed to study hot plasmas in the laboratory and natural plasmas in space. Moreover, the remarkable discoveries of the solar wind and of the Van Allen radiation belts demonstrated that plasma physics is essential to understand the space environment of the

---

This research briefing is based on the report of the Subpanel on Space and Astrophysical Plasma Physics of the *Survey of Physics* (in preparation [see App. B]), and on the report of the Committee on Solar and Space Physics of the Space Science Board entitled *Solar System Space Physics in the 1980's; a Research Strategy* (1980). This briefing was reviewed and endorsed by the Committee on Solar and Space Physics on August 1, 1984.

*Plasma physics combines concepts from electromagnetism, fluid physics, statistical mechanics, and atomic physics into a unified methodology for the study and practical use of the nonlinear collective interactions of charged particles with each other and with electric and magnetic fields. All physical conditions involving thermal energy densities exceeding 1 electron volt per atom must use plasma physics. A fluid description, magnetohydrodynamics, is used to describe the large-scale properties of plasma systems, and plasma kinetic theory is used to describe the microscopic processes that influence the large-scale properties.

earth and its effects on civilian and defense systems operating in this environment.

The two principal applications of contemporary plasma physics are to fusion research and to solar-terrestrial research—the study of the chain of physical processes that starts with the generation of the sun's magnetic field in the solar interior and links it to activity at the sun's visible surface and, ultimately, to the earth's ionosphere and atmosphere (Section 2). Our understanding of the plasma processes occurring in this interaction chain has made rapid progress in the past decade. The increasing precision of measurements, numerical modelling, and theory applied to solar-terrestrial plasma problems amounts to a revolution in technique relative to 10 years ago. The experimental diagnosis and theoretical interpretations of many space plasma processes now match in precision the best of current laboratory practice. As a result, the study of space plasmas has also become one of the most active areas of basic plasma research.

Research on solar-terrestrial plasmas is advancing our theoretical understanding of many basic plasma processes—for example, how convective turbulence generates magnetic fields, how plasma systems adjust their structure and dynamics by the process of magnetic reconnection, and how particles are accelerated to high energies. These and many other space plasma processes are nonlinear, and their study is contributing to a new branch of mathematical physics—nonlinear dynamics—which has recently led to exciting advances in many branches of physical science.

An increasingly clear perception of the functions of the links in the solar-terrestrial interaction chain has led to a design for a unified study of the solar-terrestrial system as a whole—the proposed multispacecraft International Solar-Terrestrial Physics Program (Section 3). Together with the Solar Optical Telescope, discussed in Section 2, this program would be the fundamental underpinning of solar-terrestrial plasma research for the next 10 years. We recommend that the United States commit to participating in the International Solar-Terrestrial Physics Program, by initiating funding in FY 1986.

This research briefing panel has the special responsibility to point out new opportunities. In Section 4, we will discuss the growing impact of solar-terrestrial research on astrophysical plasma physics. The plasma phenomena in the solar system have proven to be examples of processes that also occur in other stars, near neutron stars and black holes, and in galaxies. The sun and solar system have become a laboratory in which astrophysical plasma processes can be studied *in situ* and with a precision attainable nowhere else. As a result, space and astrophysical plasma physicists have begun to work closely together over the past 5 years, and a new and broad research field is beginning to develop.

The great importance of numerical computations to all plasma physics is taken up in Section 5. Such computations have already critically advanced research on fusion plasmas, and we are confident they can be equally important to solar-system and astrophysical plasma physics. We recommend early initiation of a program of large-scale numerical computations dedicated to these subjects, because detailed mathematical models clarify fundamental understanding, strengthen the interpretation of observations, and, in the case of the solar-terrestrial interaction chain, might ultimately lead to some predictive capability. Such a program would also encourage organizationally the already fruitful interactions between space and astrophysical plasma research.

2 THE SOLAR-TERRESTRIAL SYSTEM

Research in this century has revealed a chain of interactions, almost all of which involve plasma physics, that connects the generation of the sun's magnetic field, activity at the surface of the sun, the generation of the solar wind, and the dynamic behavior of the earth's magnetosphere and upper atmosphere.

The first link in the interaction chain is the sun's magnetic field. Historically, the subject of solar-terrestrial physics began with the observation that auroral and magnetic activity at earth is correlated with the 22-year cycle of the solar magnetic field. Today, no problem in solar-terrestrial research presents a greater challenge than understanding the generation and dynamics of the solar magnetic field. How does convective turbulence below the visible surface of the sun create the sun's weak magnetic field and the persistent concentrations of magnetic field that are sunspots? Recent observations of the microstructure of the sun's magnetic field challenge our understanding of solar turbulence: almost all the surface field is concentrated into tiny regions covering only 1/1000 of the surface area. The ultimate goal remains to understand how the remarkably well-ordered solar cycle arises out of such turbulent chaos.

New observational tools are at hand to help answer some of the challenging questions before us. By observing the periodic motions at the surface of the sun, the new technique of solar seismology is now revealing for the first time the structure of the solar convection zone, at the depths within the sun where the solar magnetic field is created. The Solar Optical Telescope, scheduled for a Shuttle flight in 1990, will measure photospheric and chromospheric structures with the extraordinary spatial resolution of 70 km, providing hitherto unobtainable information essential to theories of magnetohydrodynamic turbulence and solar-surface magnetic fields and dynamics. Later, this instrument is scheduled to become the principal component of the Advanced Solar Observatory, to be flown on a continuing basis on an orbiting platform.

The turbulence in the sun's outer layer heats the tenuous plasma in the solar corona to a temperature of several million degrees; however, the exact mechanism by which the plasma is heated remains uncertain. Ten years ago, *Skylab* observations revealed that the magnetic field emerging from the visible surface evolves into a complex set of open and closed coronal structures that persist for several solar rotation periods (27 days) and mysteriously resist destruction by the sun's differential rotation. The closed field regions—coronal loops—appear to channel energy into the corona and keep it there until microscopic plasma processes dissipate it. These regions are the sites of many small flares, and occasionally of a giant flare, in which magnetic energy is converted into heat and energetic particles, perhaps by a magnetic reconnection process. On the other hand, the energy flowing outward on open field lines drives a plasma wind that expands into interplanetary space. This solar wind carries hot coronal plasma and the sun's open magnetic flux past all the planets of the solar system, interacting with each to form a distinctive magnetosphere.

Our next challenge is to increase our understanding of coronal structure and plasma dynamics and of the processes regulating the flow of energy at each level of the sun's atmosphere. We hope ultimately to integrate this understanding into a unified model that accounts for coronal heating, solar flares, and the generation of the solar wind.

The solar wind is a supersonic, strongly ionized flow that carries plasma, energy, angular momentum, and magnetic field throughout the solar system, interacts with the planets and comets, and fills a three-dimensional volume of space called the heliosphere. The heliosphere terminates where the solar wind is decelerated to subsonic speeds by its interactions with interstellar matter. The solar wind is an excellent laboratory for the investigation of plasma processes of importance to basic plasma physics, solar-terrestrial research, and astrophysics. For example, fully developed magnetohydrodynamic turbulence can be conveniently studied in the solar wind but not in the laboratory. The variability of the solar corona creates a variety of structures propagating in the solar wind whose study has enriched our understanding of magnetohydrodynamics. One of these, the collisionless interplanetary shock, which triggers magnetic storms at earth, accelerates energetic particles by processes thought similar to those that accelerate galactic cosmic rays. The cosmic rays from the galaxy diffuse into the solar system by scattering from solar wind turbulence. As a result, the galactic cosmic-ray flux received at earth is largest at the minimum activity phase of the solar cycle.

Our next challenge is to synthesize our growing understanding into quantitative models that relate coronal activity to the three-dimensional behavior of the solar wind and its propagating structures. To do so, we will need to understand the microscopic processes that regulate plasma transport and the acceleration and propagation of energetic particles.

The solar wind is the plasma link between sun and earth. The solar wind does not directly hit the earth, but is largely deflected by the earth's magnetic field; the name magnetosphere refers to the region enclosing the earth in which the geomagnetic field organizes the behavior of the plasma. Within the magnetosphere, the geomagnetic field traps energetic particles to form radiation belts whose intensity is regulated by microscopic plasma turbulence. Solar-wind energy is coupled into the magnetosphere by the process of magnetic reconnection. When it occurs in the magnetopause, the boundary layer separating the solar wind from the magnetosphere, the reconnection opens part of the earth's magnetic field to the solar wind and stretches the open field into a long magnetic tail extending hundreds of earth radii downstream. A compensating reconnection process in the magnetic tail converts magnetic energy into particle energy and generates strong magnetohydrodynamic flows, which are directed away from the earth tailward of the reconnection region and towards the earth's night side earthward of the reconnection region. Although the consequences of reconnection are clear, the plasma physics of reconnection is not well understood. In the magnetopause, reconnection appears to be spatially structured and temporally variable. In the geomagnetic tail reconnection can be either steady or explosive. One thing is certain: the variability of the solar wind stimulates violent, unsteady magnetospheric flows, which lead to events called substorms, whose effects are felt in the auroral ionosphere and throughout the magnetosphere and tail.

The earth's magnetic field couples the flow in the magnetosphere and tail to the ionosphere and upper atmosphere. Nonlinear plasma processes associated with the electrical currents directed along the field lines connecting flowing magnetospheric plasma to the ionosphere accelerate electrons to about 10 keV. These electrons create the aurora when they hit the atmosphere. The same processes also heat ionospheric plasma, which then boils back out into space to modify the magnetohydrodynamic flow. Some scientists have suggested that the solar-terrestrial interaction chain reaches deep into the atmosphere to modify terrestrial weather and climate; however, the mechanisms by which this might occur have not been identified.

Thus far, solar-terrestrial research programs have concentrated on elucidating individual links in the causal chain connecting the solar wind to the magnetosphere and ionosphere. Our challenge today is to increase our understanding of each of the local plasma processes influencing the magnetosphere's structure and dynamics so that we can begin to create quantitative models that start with the solar wind's observed state and calculate the magnetosphere's and ionosphere's responses.

The most spectacular manifestation of the solar-terrestrial interaction chain is the magnetic storm. The first evidence that a large solar flare might occur is the appearance of a complex sunspot group in the sun's photosphere. Prompt electromagnetic radiation arrives at earth a few minutes after the energy in the coronal magnetic fields associated with the sunspot group is suddenly released. Energetic solar-flare protons are guided by the solar wind and magnetospheric magnetic field into the earth's polar atmosphere soon afterward. The enhanced ionospheric plasma produced by the energetic protons attenuates the radio noise received from cosmic radio sources. Several days later, a shock wave passes over earth, enveloping it in dense, hot solar-flare plasma that compresses the magnetosphere. Substorms increase in frequency and strength, expel hot plasma tailward, and inject hot plasma into the earth's equatorial inner magnetosphere to form a "ring current," which creates the geomagnetic field depression and activity that first motivated the name magnetic storm. The aurorae intensify and move to low latitudes, creating a dense highly disturbed ionospheric plasma that interferes with radio communications. Intense wind systems, sometimes of a worldwide scale, are generated in the upper atmosphere.

Many practical systems, both civilian and defense, and all our manned space endeavors must operate in the highly variable and potentially hostile plasma environment of the earth and solar system. Plasma processes in this environment also influence and even disrupt important ground-based systems over local and regional scales. Ground-based high-frequency communication systems in the earth's polar regions can be entirely blacked out by magnetic storms. Spacecraft have been electronically disabled by violent electrical discharges that occur when hot plasma from the magnetic tail envelops the spacecraft. The risk of such disasters can be reduced only by continuing attention to the effects of the plasma environment on spacecraft systems. It is clear that to work in the space environment, we must understand it.

3 THE INTERNATIONAL SOLAR-TERRESTRIAL PHYSICS PROGRAM

There is general agreement within the research community about the objectives of Solar-Terrestrial Research (Appendix A). Furthermore, there is remarkable unanimity that we can now begin the quantitative study of the solar-terrestrial system as a whole. The unanimity is expressed in the proposal currently before the administration to initiate U.S. participation in the International Solar-Terrestrial Physics Program (ISTP). Spacecraft provided by the National Aeronautics and Space Administration (NASA), the European Space Agency (ESA), and the Japanese Institute for Space and Aeronautical Sciences (ISAS) would make simultaneous, coordinated, detailed measurements of key links in the solar-terrestrial chain. The ISTP program would consist of six elements. Two spacecraft would study the sun and solar wind—ESA's Solar and Heliospheric Observatory (SOHO) and NASA's Wind spacecraft. Three would study critical regions of the earth's magnetosphere identified by previous research—POLAR (NASA), EQUATOR (NASA), and GEOTAIL (ISAS) (their names indicating the regions to be studied). ESA's multiple spacecraft CLUSTER would study small-scale plasma turbulence and boundary layers in the magnetosphere. Modern communications and data processing technology will make it possible for scientific investigators to obtain a unified set of data from all ISTP spacecraft, a prerequisite to quantitative understanding of the solar-terrestrial system as a whole.

The ISTP draws on the experience of previous coordinated programs, such as the

International Geophysical Year (1957-58) and the International Magnetospheric Study (1977-80).[*] The ISTP is a central part of a more comprehensive effort that includes other funded spacecraft, such as the International Solar-Polar Mission (NASA-ESA), the Upper Atmospheric Research Satellite (NASA), and EXOS-D (Japan). In addition, many ground-based facilities, airborne observatories, balloon and rocket campaigns, and other spacecraft can provide data for collaborative research projects. Active plasma experiments, carried out by Shuttle and other means, will enhance, and be enhanced by, the ISTP.

The United States has already provided much of the leadership in conceiving and designing ISTP, and ISTP will enable the United States to secure a position of future leadership. If it is successful, the ISTP data network, and the collaborative mode of research engendered by its existence, might well outlive the original ISTP spacecraft and shape the way that future solar-terrestrial research will be carried out.

Only a coordinated program, such as ISTP, can study the complex interaction between the parts and the whole of the solar-terrestrial system—an interaction characteristic of all large-scale plasma systems. By studying individual links in the interaction chain, ISTP will provide qualitatively new measurements critical to understanding many fundamental plasma processes—such as dynamo magnetic field generation, particle acceleration, and magnetic field reconnection—that are also important to both laboratory and astrophysical plasma research. By studying several links in the chain simultaneously, ISTP will define the contexts in which these processes occur, a prerequisite for quantitative understanding. By studying the system as a whole, ISTP will motivate and test a series of increasingly comprehensive models, whose goal is first to understand, and then to predict, the highly variable plasma environment of the sun, solar wind, and earth. More and more practical systems will have to function in this plasma environment. For all the above reasons,

> *We recommend: that the United States initiate its participation in the International Solar Terrestrial Physics Program as scheduled in FY 1986.*

4 THE IMPACT OF SOLAR-TERRESTRIAL PLASMA PHYSICS ON ASTROPHYSICS

Recent events have reaffirmed the essential unity of laboratory, solar-system, and astrophysical plasma physics that was perceived by the pioneers of the subject 50 years ago. The three branches of plasma physics are unified by a shared preoccupation with a common set of basic physical problems (Appendix B). Yet, during the past 25 years, each branch developed more or less independently, as each individually explored the opportunities of technique and application inspired by modern laboratory, space, and computational technology. A new trend is beginning to emerge. The realization is growing that solar-terrestrial research has developed problems in plasma physics that not only are fundamental to basic plasma physics but are broadly applicable to many astrophysical systems. In particular, a new field of space and astrophysical plasma physics is developing that takes as its starting point the great advances in all three branches of plasma physics during the past 25 years.

The first successes of this new, broader discipline are already in evidence. Terrestrial magnetospheric physics provided a coherent framework for understanding the magnetospheres of Mercury, Jupiter, and Saturn. The upcoming spacecraft encounters with Uranus and Neptune will further test and extend magnetospheric concepts. Since Venus is unmagnetized, studies of Venus' interaction with the solar wind will provide a foundation for investigating the plasma physics of cometary magnetospheres, which soon will be studied by spacecraft for the first time.

---

[*] Neither the solar-terrestrial interaction chain nor the role of plasma processes in it were well defined at the time of the International Geophysical Year, and the International Magnetospheric Study did not include the sun and solar wind, and so did not attempt to study the whole interaction chain.

Magnetic fields thread their way among planets, stars, galaxies, and probably even clusters and superclusters of galaxies. By transporting angular momentum, turbulent magnetic fields may regulate the accretion flow onto white dwarfs, neutron stars, and black holes. The gravitational energy released when accreting matter falls onto such compact objects may power the enormous energy outputs of galactic x-ray sources and quasars. The strong convective turbulence and buoyancy forces that develop in any gravitationally bound atmosphere will rapidly dissipate and expel magnetic fields. Hence, in nearly all astrophysical bodies, new magnetic fields must be continuously regenerated by turbulent magnetic dynamo processes. The sun provides the only accessible laboratory in which the dynamo mechanism in a highly conducting plasma can be observed and our theoretical dynamo models tested.

Plasma outflows like the solar wind are found throughout the cosmos. The *Einstein* spacecraft made the startling discovery that most stars have hot x-ray emitting coronae, implying that convective envelopes, magnetic dynamos, surface magnetic activity, and strong winds are common. Much of the interstellar medium is filled with hot, low-density plasma from blended stellar winds and supernova remnants. Since the densities, velocities, temperatures, and magnetic field strengths in the interstellar plasma are similar to those of the solar wind, solar system measurements are automatically relevant to the physics of the interstellar medium. Surrounding gas pressure collimates the winds from newly forming stars into bipolar jets; similar jets are observed in the exotic compact object, SS-433, and in radio galaxies. Relativistic winds, possibly consisting of electrons and positrons, are thought to flow away from pulsars and supermassive black holes in active galactic nuclei. The basic physics that underlies all these astrophysical winds was formulated and observationally confirmed in studies of the solar wind.

Super-energetic plasmas occur throughout the astrophysical universe. Quasars and active galactic nuclei convert prodigious amounts of energy into the relativistic plasmas that form the galactic scale radio jets. Trillion-volt electric potentials develop near the magnetic polar regions of pulsars. Radiation from the particles accelerated in these potentials is thought to trigger a pair-production cascade that fills the surrounding magnetosphere with an electron-positron plasma and powers a highly relativistic wind. In our galaxy, supernova shocks accelerate the cosmic rays up to a total energy density that is comparable with the thermal and magnetic energy densities of the interstellar medium. The plasma processes responsible for particle acceleration have long been studied in the solar-terrestrial context—in planetary magnetospheres, at planetary bow shocks and interplanetary shocks, and in solar flares. The acceleration models derived from these observations guide and test the underlying plasma physics of their astrophysical counterparts. For example, although it is not possible to detect the plasma processes responsible for particle acceleration in supernova shocks, theories of cosmic ray acceleration can be tested by measurements of solar system shocks together with measurements of galactic cosmic ray energy spectra and composition.

The challenge for the next decade is becoming clear: space and astrophysical plasma physicists must work side by side to formulate quantitative models of solar system and astrophysical phenomena in which magnetohydrodynamics and plasma physics are essential for the interpretation of observations, recognizing that most of the universe is in the plasma state.

5 LARGE-SCALE NUMERICAL MODELS

Plasma physics was a pioneer in the successful utilization of large-scale computations for fluid, magnetohydrodynamic, hybrid, and kinetic models. Most of the progress since the late 1960s has been in fusion research and nuclear weapons phenomenology. The establishment of a computational facility dedicated to magnetic fusion significantly advanced understanding of magnetic confinement systems and of fusion and basic plasma processes. The establishment, in 1979, of NASA's highly successful Solar-Terrestrial Theory Program made numerical models and simulations regular tools in solar system plasma research and has prepared that research community for the next, more advanced stage.

In the past decade, our research on solar-terrestrial plasmas has achieved a measure of quantitative understanding through the use of numerical models. The study of plasmas beyond the solar system has developed more slowly than space plasma physics for a fundamental reason: the microscopic plasma processes that regulate the behavior of astrophysical systems cannot be observed directly, as can those in the solar system and in the laboratory. Now, however, the modern theoretical and computational techniques developed to understand laboratory and space measurements have opened the door to modelling of the plasmas in the still larger and more exotic environments of astrophysics. Since the underlying astrophysical plasma processes cannot be directly detected, we believe that the best strategy will be to create numerical models at the large-scale system level that postulate small-scale plasma and radiation processes and to iterate between the system and process levels until quantitative agreement with observation is achieved. It is our perception that, for the first time, the current level of development of numerical technology, theory, and observations gives such a strategy a significant chance of success.

The continued development of numerical technology will advance many branches of science. We foresee that many problems of a scale that requires today's national computing facilities will soon be addressable by local university and laboratory facilities. This will only increase the importance of numerical modelling to our subjects. Nonetheless, we believe that the leading research on many solar system and astrophysical plasma models will continue to be done on the most advanced computing facilities existing at any given time, because these models involve a complex interplay between large- and small-scale processes.

Thus far, the responsibility for the maintenance and advancement of state-of-the-art computing facilities has been a national one, because it is beyond the capability of single institutions and because a national scope provides an adequate pool of users. America's existing advanced computational facilities, devoted to defense, fusion research, and meteorology, have been used on a piecemeal basis for solar system and astrophysical plasma research. These busy facilities do not have space physics and astrophysics as an institutional objective, and researchers in these fields must make individual agreements to secure access to advanced computing. In some cases, American researchers have had to journey to Europe or Japan in order to perform large-scale computations.

To enhance the research of the International Solar-Terrestrial Physics Program, and to promote fruitful interchanges between solar system and astrophysical plasma physics,

We recommend: a national computational program dedicated to basic plasma physics, space *physics, and astrophysics, which will provide and maintain state-of-the-art technology appropriate to large-scale theoretical models and simulations. Such a program should ensure ready access to advanced computing on the basis of peer review.*

Many problems central to solar-system and astrophysical plasma research are ready for advanced numerical modelling. If these problems are combined with others in hydrodynamics and general astrophysics that should be included in the program, the increased scientific demand, and more importantly, the scientific payoff, would justify the dedicated effort we propose. The success of numerical computations applied to laboratory plasma problems gives us confidence that a computational program would significantly advance space and astrophysical plasma research. Solar-terrestrial research would be the principal initial contributor to, and beneficiary of, such a program. Solar-terrestrial models that successfully meet the test of detailed measurements of both large- and small-scale processes would substantially increase our confidence in models of more distant astrophysical systems.

Appendix A

The Goals of Solar-Terrestrial Physics

The research goals of solar-terrestrial physics have been succinctly defined by the Space Science Board in *Solar-System Space Physics, a Research Strategy* (National Academy of Sciences, 1980).

Two basic principles guided the formulation of these goals:

- The objectives of solar-system space research are to understand the physics of the sun; the heliosphere; and the magnetosphere, ionosphere, and upper atmosphere of the earth, other planets, and comets.
- Studies of the interactive processes that generate solar variation and link it to the earth should be emphasized, because they reveal basic physical mechanisms and have useful applications.

To understand better all the processes linking the solar interior to the corona, we need to study the following:

- The sun's global circulation, how it reflects interior dynamics, could modify luminosity, and is related to the solar cycle.
- The interactions of solar plasma with strong magnetic fields-active regions, sunspots, and fine-scale magnetic knots—and how they cause the release of solar-flare energy to the heliosphere.
- The solar corona's energy sources and the physics of its large-scale weak magnetic field.

To understand better the transport of energy, momentum, energetic particles, plasma, and magnetic field through interplanetary space, we need to study the following:

- First and foremost, the coronal processes that govern the generation, structure, and variability of the solar wind.
- The three-dimensional properties of the solar wind and heliosphere.
- The plasma processes that regulate solar-wind transport and accelerate energetic particles throughout the heliosphere.

To understand better the time-dependent interaction between the solar wind and earth, we need to study the following:

- The transport of energy, momentum, plasma, and magnetic and electric fields across the magnetopause.
- The storage and release of energy in the earth's magnetic tail.
- The origin and fate of the plasma(s) within the magnetosphere.
- How the earth's magnetosphere, ionosphere, and atmosphere interact.

To understand better the entire upper atmosphere as one dynamic, radiating and chemically active fluid, we should study the following:

- The radiant energy balance, chemistry, and dynamics of the mesosphere and stratosphere and their interactions with atmospheric layers above and below.
- The worldwide effects of the magnetosphere's interaction with the polar thermosphere and mesosphere and the role of electric fields in the earth's atmosphere and space environment.
- The effects of variable photon and energetic particle fluxes on the thermosphere and on chemically active minor constituents of the mesophere and stratosphere.

To understand better the effects of the solar cycle, solar activity, and solar-wind disturbances upon earth, we need to:

- Provide to the extent possible simultaneous measurements on many links in the chain of interactions linking solar perturbations to their terrestrial response.
- Create and test increasingly comprehensive quantitative models of these processes.

To clarify the possible solar-terrestrial influence on earth's weather and climate, we need to:

- Determine if variations in solar luminosity and spectral irradiance sufficient to modify weather and climate exist.
- Ascertain whether any processes involving solar and magnetospheric variability can cause measurable changes in the earth's lower atmosphere.
- Strengthen correlation studies of solar-terrestrial, climatological, and meteorological data.

Appendix B

The Unity of Laboratory, Solar System, and Astrophysical Plasma Physics

Although the experimental techniques we use to study laboratory, solar system, and astrophysical plasmas differ, our research on them is unified by the use of plasma concepts. The Committee on Space Physics (National Academy of Sciences, 1984) and the Subpanel on Space and Astrophysical Plasma Physics of the *Survey of Physics* (National Academy of Sciences, in preparation) have concluded that these disciplines share a common set of basic physical problems. These include:

1. magnetic field reconnection;
2. the interaction of turbulence with magnetic fields;
3. the behavior of large-scale plasma flows and their interactions with magnetic and gravitational fields;
4. the acceleration of energetic particles;
5. particle confinement and transport;
6. collisionless shocks;
7. beam-plasma interactions and the generation of electromagnetic radiation; and
8. collective interactions between neutral gases and plasmas.

The fact that such problems emerge from a variety of contexts demonstrates their general significance and suggests that their solution will find further applicability in contexts we cannot imagine today. The existence of such general problems provides a basis upon which a network of common interests, personal interactions, and a common discipline are being built.

Report of the Research Briefing Panel on Selected Opportunities in Physics

Research Briefing Panel on Selected Opportunities in Physics

Hans Frauenfelder (*Co-chairman*), University of Illinois, Urbana, Ill.

Mildred S. Dresselhaus (*Co-chairman*), Massachusetts Institute of Technology, Cambridge, Mass.

William F. Brinkman, Sandia National Laboratory, Albuquerque, N. Mex.

Praveen Chaudhari, International Business Machines Corporation, Yorktown Heights, N.Y.

Joseph Cerny, Lawrence Berkeley Laboratory, Berkeley, Calif.

John M. Dawson, University of California at Los Angeles, Los Angeles, Calif.

Harold P. Furth, Princeton University, Princeton, N.J.

Daniel Kleppner, Massachusetts Institute of Technology, Cambridge, Mass.

Alexei A. Maradudin, University of California at Irvine, Irvine, Calif.

Francis M. Pipkin, Harvard University, Cambridge, Mass.

John Schiffer, Argonne National Laboratory, Argonne, Ill.

Watt W. Webb, Cornell University, Ithaca, N.Y.

David Wilkinson, Princeton University, Princeton, N.J.

Staff

Donald C. Shapero, *Staff Director*, Board on Physics and Astronomy, National Research Council

Ellen Williams, *Rapporteur*, University of Maryland, College Park, Md.

Helene Patterson, *Staff Assistant*, Board on Physics and Astronomy, National Research Council

Florence Wong, *Secretary*, Board on Physics and Astronomy, National Research Council

Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

The panel gratefully acknowledges assistance in preparing its report from Gordon Baym, Robert B. Palmer, Bruce Partridge, David Schramm, Lee S. Schroeder, Andrew M. Sessler, Charles P. Slichter, and Robert Wagoner. Thanks are also due to participants in a planning meeting to select the topics treated in the Physics Research Briefing; those not named above include Richard Brewer, IBM; Dale Compton, Ford Motor Company; Robert Park, University of Maryland and American Physical Society/Washington; and Robert Richardson, Cornell University.

Report of the Research Briefing Panel on Selected Opportunities in Physics

INTRODUCTION

Physics has been rich in discoveries of fundamental laws of nature. These discoveries have influenced other fields of science from mathematics to biology and medicine and have spawned entire industries. Maxwell's electromagnetic theory, for example, forms the basis for understanding electromagnetic phenomena from waves to plasmas and underlies radio, television, radar, and our industrial power networks and modern communication systems. Quantum mechanics provides a framework for portraying physical reality and underlies all of the natural sciences. It has made possible the invention of the transistor, many of the amazing devices of modern solid state electronics, and the laser. Lasers, in turn, are having a major impact on many areas of science. They are also playing a rapidly expanding role in medicine and broad areas of technology, including communications, manufacturing, and national defense.

The discovery of x rays led to the largest single advance in the history of medical diagnostics. The recently created NMR body imaging technique may mark a comparable advance. From the study of electrical noise in the atmosphere, radioastronomy was born, which in turn provided a dazzling new portrait of the universe. Investigations of the properties of nuclei led to the creation of the nuclear power industry. Numerous other examples could be cited. The benefits to society from basic research in physics have been incalculable.

Recent fundamental advances in physics demonstrate that physics is still in a golden age and has never been more vigorous or more productively interactive with other fields. Progress in physics over the last decade has been remarkable. Puzzles that seemed to present insuperable challenges at the beginning of the 1970s have yielded to powerful and elegant theoretical and experimental techniques. The new insights and accomplishments have not only brought greater unity to the various branches of physics but have also strengthened the ties of physics to other areas of science and opened a vast array of new opportunities. Every part of physics has participated in the advance, as the forthcoming report of the Physics Survey Committee will make clear.

The present report focuses on six areas of special opportunity in physics. In order to find areas in which incremental funding may lead to major advances, a planning group (consisting of the Briefing Panel and others listed above) was convened. From twenty suggestions, the following six were chosen: (1) physics at the laser-atomic frontier, (2) relativistic plasma waves, (3) physical properties of deliberately-structured materials, (4) biomolecular dynamics and intercellular cooperativity, (5) cosmology, and (6) nuclear matter under extreme conditions. Large national facilities and well-established programs were explicitly excluded.

The six areas chosen promise to yield fundamental results of great interest. Many of the areas are likely to advance technology and to produce results with impact on the nation's industries; many can be expected to contribute to national security.

All six areas cut across lines of narrow specialization and link different fields. Progress in any one area can thus be expected to stimulate progress in other fields. Cosmology, for instance, connects physics to astronomy and astrophysics. Biological physics interacts strongly with chemistry, biochemistry, biology, medicine, and pharmacology. Deliberately-structured materials relate fundamental aspects of physics to many fields of technology.

1. PHYSICS AT THE LASER-ATOMIC FRONTIER

The laser continues to have a revolutionary impact in the field from which it emerged — atomic physics. The laser-atomic frontier comprises research opportunities generated by joining laser methods with new techniques in atomic and molecular physics.

Opportunities at the laser-atomic frontier include research with trapped particles; femtosecond (10^{-15} sec) spectroscopy and other new forms of spectroscopy; the study of previously inaccessible atomic, molecular, and ionic species; and the development of novel light sources. Advances in these areas will lead to new tests of basic theory and deepen our understanding of the structure of matter at the atomic and molecular level, including the transfer of energy and the nature of chemical reactions. In addition to high scientific interest, many of these topics are likely to yield technological advances. The opportunities selected include the following.

Trapped Particles

Using laser light, it is possible to cool to the millikelvin region ions confined in electromagnetic traps and also to cool and trap neutral atoms. In conjunction with recently developed ultrastable tunable lasers, these developments open the way to major advances in high-precision measurements: frequencies to parts in 10^{16}, masses to parts in 10^{11}, new tests of the isotropy of space, and studies of collective motion in plasmas and gases. Applications include advanced atomic clocks, optical frequency standards, metrology, and communications.

New Spectroscopies

Breakthroughs in the generation of femtosecond light pulses make it possible to take "snapshots" of atoms as they collide or molecules as they react, and to observe fast adsorption and relaxation on surfaces. Femtosecond spectroscopy has applications to fast electron circuitry and high-speed instrumentation. Another new type of spectroscopy is carried out with relativistic particle beams. Using the Doppler effect to shift highly stable laser radiation from the visible into the ultraviolet opens the way to new areas of high-resolution spectroscopy of atoms, ions, and molecules. Scientific objectives include studies of relativistic and quantum electrodynamic effects in highly charged ions, high resolution ultraviolet spectroscopy of atoms and molecules, and new types of photoejection measurements.

The Creation of New Species

Lasers can produce new species as well as species that have been inaccessible, such as multiply excited atoms, molecular ions, and clusters. Problems such as correlated electron motion, the evolution of matter from single atoms to the condensed state, and catalysis can be investigated. There are many applications to other sciences and to industrial problems. Studies of ionic species can lead to the design of more efficient combustion engines; catalysis is of enormous importance to chemical processing.

MATTER IN INTENSE FIELDS

Laser methods open the way to the study of matter under extraordinary conditions, such as in intense electric and magnetic fields and powerful radiation fields where nonlinear phenomena and multiphoton processes occur. Studies of multiphoton processes made possible by high-power lasers are expected to reveal new aspects of the interaction of photons with matter, and they may open the way to novel types of photochemical processing and isotope separation.

NEW LIGHT SOURCES

The creation of the excimer laser and the generation of extreme ultraviolet light in supersonic atomic beams are two recent examples of light sources from research at the laser-atomic frontier. Other new sources can be expected, with wavelengths from the infrared to the soft x-ray region. Such sources would be valuable to wide areas of atomic physics, chemistry, condensed matter physics, and biology. Using coherent generation of pulsed laser light, for example, it may be possible to provide very short pulses in the soft x-ray region. This technique may provide a laboratory-based alternative to synchrotron radiation light sources for many applications.

2. RELATIVISTIC PLASMA WAVES

Recently an exciting area of plasma physics has emerged — that of relativistic plasma waves in which the electrons or both electrons and ions have relativistic velocities. Exploration of relativistic plasma waves is expected to elucidate recently discovered exotic astrophysical objects and the acceleration of particles to create cosmic rays. An understanding of relativistic collective electromagnetic phenomena may lead to novel particle accelerators and radiation generators. We sketch below some of the specific ideas that need investigation.

PARTICLE ACCELERATORS

The next generation of high-energy accelerators is pushing existing technology close to the limit. The long-range future of particle and nuclear physics may well depend on the development of new acceleration techniques. In addition, new accelerators in the range of 10 MeV to 10 GeV may find application in physics, other sciences, industry, and medicine.

Plasma Accelerators

Space-charge waves in plasmas can produce electric fields that are orders of magnitude larger than can be generated by any other means, with the possible exception of a laser beam focused in a vacuum. These fields can be larger than the fields that bind electrons to atoms (10^9 volts/cm). Such fields can only be maintained in a plasma because they instantly ionize ordinary materials. In a beat-wave accelerator, space-charge waves are generated by sending through the plasma two intense laser beams with a frequency difference equal to the plasma frequency. The combination of high electric fields and phase velocities close to that of light makes such waves good candidates for the acceleration of particles to high energy.

Grating Accelerators

Another novel method uses a grating of droplets illuminated by a powerful short-pulse laser. The resulting plasma grating generates fields estimated to be of several hundred million volts per cm.

Free-Electron Laser Accelerators

Another possibility for constructing an accelerator is to operate a free electron laser "backwards" in an accelerating mode. A further possibility is to operate a free electron laser "forward" and use the resulting radiation to accelerate particles in a conventional linear accelerator. Both of these concepts pose interesting problems and will require developments in current technology.

Laser Development

All of these accelerating schemes require terawatt lasers with good optical quality and very short pulses (shorter than 1 picosecond). Progress with these approaches could be accelerated by a laser development program aimed toward these requirements.

Generation of Electromagnetic Radiation

The inverse of particle acceleration—radiation by energetic particles in plasmas—is a challenging subject of scientific and potentially practical interest. Bursts of radiation that accompany solar activity and are also emitted from a wide variety of astrophysical objects are poorly understood. Recent observations of radiation from pulsars show that the pulses have γ-ray energies up to 10^{15} eV. Such energies imply that the acceleration process occurs over relatively short distances (of the order of 10^3 km) with accelerating fields larger than 10^7 volts/cm. Relativistic waves may be involved. Closer to home, bursts of energetic particles are generated by solar flares; here again strong electric fields must exist in the plasma. The theoretical and computer modeling tools are now reaching the point where progress in understanding these processes can be made.

In the laboratory, copious radiation, roughly up to the thirtieth harmonic of the plasma frequency, has been observed when relativistic electron beams are injected into plasmas. The generation of electromagnetic radiation by means of free-electron lasers is actively pursued at a number of laboratories. Within the last year, free-electron lasers have started operation—two in the microwave, three in the infrared, and one in the visible range. Further configurations especially suited to different parts of the spectrum could be developed. Moreover, theory predicts that energetic electrons streaming through plasmas with strong density fluctuations should yield radiation similar to that produced in present free electron lasers where rippled magnetic fields are used instead of density ripples. At present there are virtually no controlled experiments to investigate these phenomena. The computational tools to address these problems exist, but the resources for using these tools are generally not available to university researchers.

3. DELIBERATELY-STRUCTURED MATERIALS

Condensed-matter science has a long and continuing history of contributions to basic understanding of physical phenomena. General concepts from this field are bringing greater intellectual unity to physics. Ideas from condensed-matter physics and many-body theory are playing a role in such diverse fields as nuclear physics, astrophysics, and theoretical methods of particle physics. While condensed-matter physics is intellectually a rich and fertile field of inquiry for the basic questions of science, it is also playing a vital role in advanced technology from microelectronics and communications to energy systems.

From the vast range of activities in condensed-matter physics, we emphasize only one area of special opportunity: deliberately-structured materials. By deliberately-structured materials, we mean materials designed by man to have particular structures. Their fabrication may involve controlling on an atomic length scale the compositional or structural arrangements of atoms in one or two dimensions. The resulting materials possess remarkable properties not otherwise found in nature—properties that, in many instances, present major scientific puzzles.

Not only are these materials likely to raise fundamental questions of physics, but their study will also reinforce the technological base of the nation's most important device-related industries. Conversely, technologies that have led to advances in the fabrication of very-large-scale integrated circuits can be harnessed to create deliberately-structured materials.

Three research areas that are currently of great scientific interest and that are especially likely to advance through the use of deliberately-structured materials are: (1) studies of surfaces and interfaces between specifically designed microscopic structures; (2) investigations of "disorder" in matter that might be understood more deeply through tailoring of its composition or geometrical structure; and (3) studies of novel collective phenomena in "exotic materials."

Surfaces, Interfaces, and Thin Films

Beginning with an atomically clean and crystallographically defined surface, it is possible to prepare a two-dimensional film only 1 atom thick by deposition of a second material. The film can be atomically smooth or contain crystallographic discontinuities, depending on the structure of the substrate. Its structure can be further modified by pattern etching on a scale finer than 100 Å. Deposition of additional layers of dopants, or of metals, semiconductors, superconductors, or insulators, leads to atomically sharp interfaces of defined structure and composition. The physics of these low-dimensional systems is fundamentally interesting. Furthermore, studies of the electronic, vibrational, magnetic, and transport properties of interfaces are applicable to the technological needs of the computer and energy industries.

Disorder in Matter

There are many approaches to creating materials with controlled disorder. One method is the random dispersion of very small particles of one material in a matrix of a second. Another is the fabrication of alternating amorphous layers, resulting in two-dimensional disorder within the plane of the layers but one-dimensional order perpendicular to the plane due to the stacking of the layers. Spin glasses present a totally different form of disorder in which the atoms may be compositionally ordered but the spin-spin interactions are random. Metastable (i.e., nonequilibrium) disordered structures represent a fourth class of deliberately-disordered materials. All these materials have properties that are intrinsically interesting and important. In addition, insight into the nature of controlled disorder may provide a basis for understanding technologically important disordered materials such as polymers, glasses, and composite materials and may even lead to a better understanding of problems in computation, optimization, and neurobiology.

Exotic Materials

The structure of materials can be arranged in many different ways to produce unique physical properties. For instance, it is possible to deliberately construct ordered crystals in arrays of alternating layers called superlattices. The properties of superlattices can be modified in desirable ways by tailoring the compositional and impurity profiles of the constituent layers. Many of the traditional methods of condensed-matter theory must be extended to account for the small size or the lower dimensionality of such systems. New phenomena, undreamt of a short time ago, have been discovered in deliberately-structured materials. Examples include quantization and fractional quantization of the Hall effect. The same class of materials that is actively used to investigate the quantum Hall effect is also being examined by researchers worldwide for new applications in the computer and communication industries.

Needs of the Field

The overwhelming characteristic of this field is its dependence on a strong infrastructure to support a range of activities from materials preparation and characterization to study of the nature and causes of unusual materials properties. The infrastructure, involving skilled technicians and complex, state-of-the-art apparatus and instruments, does not exist in most universities; and the institutional mechanisms to put the structure in place at universities are often lacking. The intellectual concepts are there to be exploited if the university infrastructure can be enhanced.

4. BIOMOLECULAR DYNAMICS AND INTERCELLULAR COOPERATIVITY

The magnificent complexity of life that we study as biology ultimately reflects underlying physical principles. At each level—biomolecules, supramolecular structures, cells, and organisms—the behavior of the system is governed by physical laws. The detailed description of biological phenomena in terms of physical laws promises to yield rich dividends both in terms of understanding life more deeply and in exploring the physics of complex systems. Research on biological systems connects physics to chemistry, biochemistry, biology, and, ultimately, pharmacology and medicine. The systems under investigation are, of course, the same in all these fields, but the approaches to them are different. The physicist selects the simplest available system, performs quantitative measurements, constructs a model, and attempts to describe the observed behavior fully in terms of known physical laws. The following examples, at different levels of complexity, indicate where this approach is beginning to bear fruit.

Biomolecular Dynamics

Proteins and nucleic acids, the building blocks of life, were called "aperiodic crystals" by Schrödinger. Seen by standard x-ray diffraction, they indeed appear to be like crystals with linear dimensions of a few nanometers. In recent years, however, a far more exciting picture has emerged. Biomolecules are dynamic systems that represent a state of matter different from solids, liquids, and gases. They show many similarities to glasses and spin glasses but are far more sophisticated. The sophistication is not surprising; each biomolecule has undergone more than 10^9 years of research and development and is built for a particular function. A protein, for instance, must have precise mechanical, electrical, thermal, and chemical properties. These properties are not spatially homogeneous but rather vary throughout the protein volume. Through genetic engineering, proteins with atomic modifications in desired places can be produced. The entire field of physical studies of biomolecules is at its very beginning, but the importance and the potential of such studies are increasingly evident. Biomolecules are excellent laboratories for the investigation of basic physical laws. We can study elementary excitations, motions, transport phenomena, and elementary reaction steps. Biomolecular physics overlaps the physics of glasses and spin glasses and complements the chemistry of reactivity, catalysis, and life processes. With increasing knowledge of the relation between structure and function, it will become possible to design and synthesize biomolecules with desired properties for uses in biological research, medicine, and pharmacology. It is also likely that an understanding of the relation between structure and function of biomolecules, together with genetic engineering, will produce spectacular results in other areas. Biosensors have the potential of improving equipment from medicine to information transfer. New biomaterials, such as biomolecular catalysts and artificial photosynthetic systems, may find uses in areas ranging from agriculture to space.

TRANSMEMBRANE SIGNALLING

Transmembrane signalling is conspicuous in brain, nerve, and muscle tissue. Its molecular basis is now directly accessible through a new biophysical technology: recording of the stochastic electrical conductances of individual molecular channels isolated in very small patches of cell membranes or reconstituted into model membranes. Switching of transmembrane currents through molecular channels is controlled by neurotransmitters, messenger molecules, or membrane potential changes. Hundreds of distinguishable membrane channels and nearly as many neurotransmitters have been discovered. How are channel protein conductances switched by binding of a neurotransmitter? How is ion selectivity built into the protein structure? These profound questions and the biological behavior and disorders associated with them appear susceptible to solution with the help of a combination of gene manipulation and biophysical measurements.

MOLECULAR BASIS OF INFORMATION STORAGE

Elegant experiments on the organization of elementary learning and memory in simple animals at the cellular and molecular level imply a variety of basic signalling processes. The molecular basis of memory is, however, still not understood. The combination of advanced biophysical techniques, such as sensitive optical and electron microscopy probes, with hybridoma techniques for producing monoclonal antibodies, offers hope of elucidating the molecular features of information storage and signal transmission. Prospects are good for understanding the molecular mechanisms of auditory transduction and tone discrimination and the mechanism of visual signal transmission from the photoreceptor. A new family of sensors may develop from these biophysical experiments with some hope of revolutionizing the tools of clinical medicine.

INTERCELLULAR COOPERATIVE PROCESSES

Recent developments at the interface between physics and mathematics have generated new analytical concepts dealing with discrete and nonlinear systems. These systems often display regimes of solution characterized by solitons, limit cycles, disorder, and chaos. These concepts represent many intercellular cooperative processes and have provided understanding of several systems, such as the cooperative heart muscle cell contraction in the heart beat and its failure in ventricular fibrillation.

Physicists have generated new conceptual models of distributed memory. These models, while speculative at present, may, if successful, find ultimate applications in robotics and computer architecture. The theory reproduces many of the foibles and perhaps some of the strengths of the human memory.

5. COSMOLOGY

Few areas of science enjoy as wide an appeal to scientists and laypersons as cosmology. Fascination with the origin and fate of the universe is deeply imbedded in the human intellect, and speculations about these ultimate questions have stimulated some of the most profound thinking in modern science. As in all other scientific disciplines, the basis of cosmology lies in experimental observations. Accurate measurements of the abundances of light elements, for example, fit the specific predictions of a "hot big-bang" model for our universe. Light nuclei were formed when the universe was a few minutes old and had a temperature of 10^9 K—a fantastic extrapolation of known physics. The same hot big-bang model predicted the 3 K cosmic radiation a decade before its discovery.

Cosmology and Particle Physics

Though uniquely successful in explaining these fundamental observations, the hot big-bang model by itself fails to explain others.

- Why is our universe so asymmetric in matter and antimatter?
- Why is the ratio of baryons to photons in the universe $\sim 3 \times 10^{-10}$?
- Why is the kinetic energy in universal expansion so nearly equal to the universe's gravitational potential energy at the current epoch?
- How did the universe reach the same temperature to a part in 10^4 at points not causally connected?

Two recent ideas from particle physics may answer these fundamental questions and still fit into the basic hot big-bang picture. Grand Unified Theories (GUTs) when applied in the very early universe (age $\sim 10^{-35}$ sec) supply answers to the first two questions. Thus, laboratory tests of GUTs (e.g., by means of proton-decay experiments) may also resolve important cosmological problems. The last two questions above may be answered by a model in which the universe expanded exponentially for a brief time early in its history. This "inflation" is supposed to have been driven by the vacuum energy of a scalar field perhaps related to the Higgs particle that is required by the new theories of electromagnetic and weak interactions. We can, of course, turn the arguments around. Already cosmological data are constraining and guiding GUTs, making the early universe effectively a particle physics laboratory.

What Lies Ahead?

The long-awaited evidence to describe the fate of the universe could come from either of two lines of attack. One approach attempts to measure directly the deceleration of the universe over the past few billion years. The other approach looks for "invisible" mass density that might eventually reverse the current universal expansion, leading to a hot recollapse. Massive neutrinos, axions, and other exotic particles are under active study in this context by particle physicists and cosmologists. Primordial nucleosynthesis has recently advanced a strong argument against the existence of a large component of dark matter in baryonic form; thus, the search for an exotic mass component has intensified.

Another area of current excitement is the origin and development of large-scale structure—galaxies, clusters of galaxies, etc. The structure seen today is assumed to grow from density perturbations in the early universe. However, observational limits on fluctuations in the thermal background radiation at small angular scale severely constrain models for the origin of structure. Most models that lead to the observed matter clumping also predict fluctuations of the radiation exceeding observed limits. The standard picture must be modified and extended.

Cosmology is at present ripe with opportunity and certain to be full of surprises. The combination of the space program with major advances in astronomical instrumentation is rapidly pushing our horizons deeper into the universe across a wide range of the electromagnetic spectrum. Theory, drawing on all areas of physics, is creating a plausible and testable cosmological model rich with beautiful physics and unanswered questions.

6. NUCLEAR MATTER UNDER EXTREME CONDITIONS

Until recently the study of nuclear matter was restricted to normal values of density and low temperature. With the acceleration of heavy nuclei to relativistic energies, an important new dimension has been added to these studies, namely, the ability to deliver unprecedented amounts of energy and momentum into the volume of the colliding nuclei, a volume large relative to that of a nucleon. With this tool a new era for studying nuclear matter opens up; for the first time we can test the response of such matter to extreme conditions of energy density and compression.

SCIENTIFIC OPPORTUNITIES

Central nucleus-nucleus collisions at high energies open up new avenues that cross the traditional boundaries of particle, nuclear, and astrophysics. The broad goals of such a program — to explore the cooperative behavior in extended hadronic systems with many interacting degrees of freedom — are complementary to those of particle physics in the same sense that condensed-matter physics, with its varied and often unexpected phenomena, transcends the underlying atomic physics. The salient difference is that nuclear studies can themselves shed light on the incompletely understood mechanism of confinement in the basic quantum chromodynamic theory.

Foremost among the opportunities is the possibility of producing a new state of matter called the quark-gluon plasma. Through numerical lattice-gauge calculations, quantum chromodynamics predicts that confined hadronic matter, in which quarks are bound together to form baryons and mesons within the nucleus, will cease to exist at sufficiently high values of compression or energy density. A new phase of matter appears in which the quarks and gluons are freed from the confines of the hadrons. Current theoretical estimates and extrapolations of existing data indicate that the phase transition can be explored in at least two distinct domains accessible in heavy-ion collisions. The domain of *maximum nuclear compression* is a baryon-rich environment expected to occur at energies of a few GeV/nucleon in the center-of-mass frame. In this domain, the heaviest nuclei can still stop each other completely. Densities up to 10 times that of normal nuclei should occur; such densities may be achieved through enhancement of existing accelerators. The domain of *high-energy density* is expected to occur in the central rapidity region left after the colliding nuclei pass through each other at energies of several tens of GeV/nucleon in the center-of-mass frame. Exploration of this domain, expected to be meson-rich with energy densities in excess of 10 times those of ordinary nuclei, awaits construction of future accelerators.

Colliding very heavy nuclei at relativistic energies allows one to study new states of matter and to map out the equation of state of highly compressed/excited nuclear matter, while providing the first indications of the dynamical behavior of extended regions of matter under the most extreme baryon and energy densities accessible in the laboratory.

Suitable facilities will soon be available for accelerating light ions to the 15-225 GeV/nucleon range. These new capabilities are in the region expected to yield maximum nuclear compression and to provide an important extension of our studies of nuclear matter under extreme conditions.

By their very nature, these complex experiments present a major challenge to experimentalists. Detector requirements are similar to those found in particle physics but with a high degree of detector segmentation and enhanced particle identification because of the large numbers of particles expected. Research and development efforts on these detectors are needed so that we will be adequately prepared to measure the complex event structures expected in these interactions.

In nuclear theory, substantial efforts are required to produce theoretical benchmarks for comparison with experimental results. Examples include large Monte Carlo intranuclear cascade calculations and lattice-gauge calculations with inclusion of finite quark masses and finite baryon densities.

IMPACT ON OTHER AREAS

In addition to potential technological spinoffs from the development of new detectors, research on central high-energy nucleus-

nucleus collisions links nuclear physics to other fields. Information gained about the transition from the quark-gluon plasma to hadronic matter is important to cosmologists for understanding how the present state and distribution of matter in the universe evolved. Data from the regime of maximum compression are required to understand the state of matter in the deep interior of neutron stars. There is no other way such information on the high-density equation of state and other properties of matter vital for condensed-matter astrophysics can be obtained in the laboratory. These data will facilitate the interpretation of observations from future orbiting x-ray telescopes such as the AXAF and will help to unravel questions about the birth and evolution of neutron stars. Particle physicists have been studying the short-range behavior of quantum chromodynamics and processes of quark confinement; high-energy nuclear collisions will allow us for the first time to investigate the long-range behavior of this theory and the processes of deconfinement.

CONCLUSIONS

The six areas selected in the present report comprise only a small part of physics. But they exhibit clearly the characteristics of physics—enormous diversity, the search for fundamental laws, strong connections to many other sciences, and technological and industrial applications. Each of the six areas is at a stage where incremental funding may produce major progress. The reasons for the high leverage vary from one area to another, but in each area, there could be rapid progress through meeting a need either for seed funding to promote new research or for new instrumentation, especially medium-cost instrumentation for university research laboratories. Several of the subjects—relativistic plasma accelerators, cosmology, and parts of biophysics—do not have established mechanisms for support and review.

The new areas within nuclear physics, biophysics, and cosmology described above are each in their infancy; exploiting the opportunities to do new forefront physics addressing fundamental questions about matter, life, and the universe could be greatly accelerated by a modest investment of seed money.

Each of the new areas within atomic physics, plasma physics, and condensed-matter physics described above has a need for medium-cost instrumentation that is difficult for university laboratories to obtain within existing funding frameworks. An increment of funding targeted to the instrumentation problem in these areas could unleash untapped scientific productivity in areas with potential applications to technology and national security.

Research Briefings 1985

National Academy Press 2101 Constitution Avenue, NW Washington, DC 20418

NOTICE: The National Academy of Sciences was established in 1863 by Act of Congress as a private, nonprofit, self-governing membership corporation for the furtherance of science and technology for the general welfare. The terms of its charter require the National Academy of Sciences to advise the federal government upon request within its fields of competence. Under this corporate charter, the National Academy of Engineering and the Institute of Medicine were established in 1964 and 1970, respectively.

The Committee on Science, Engineering, and Public Policy is a joint committee of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. It includes members of the councils of all three bodies.

This work was supported by the National Science Foundation under Contract PSP791304.

Library of Congress Catalog Card Number 85-61729

International Standard Book Number 0-309-03585-6

Copyright © 1985 by the National Academy of Sciences

No part of this book may be reproduced by any mechanical, photographic, or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted, or otherwise copied for public or private use, without written permission from the publisher, except for purposes of official use by the United States government.

Printed in the United States of America

Committee on Science, Engineering, and Public Policy

LEON T. SILVER, William M. Keck Foundation Professor of Geology, Division of Geological and Planetary Sciences, California Institute of Technology (*Chairman*)

*LINDA H. AIKEN, Vice President for Research, The Robert Wood Johnson Foundation, Princeton, New Jersey

RICHARD C. ATKINSON, Chancellor, University of California, San Diego

JACOB BIGELEISEN, Leading Professor, Department of Chemistry, State University of New York, Stony Brook

FLOYD E. BLOOM, Director and Member, Division of Pre-Clinical Neuroscience and Endocrinology, Scripps Clinic and Research Foundation, La Jolla, California

W. DALE COMPTON, Vice President, Research, Ford Motor Company, Dearborn, Michigan

EMILIO Q. DADDARIO, Wilkes, Artis, Hedrick and Lane, Attorneys at Law, Washington, D.C.

GERALD P. DINNEEN, Vice President, Science and Technology, Honeywell, Inc., Minneapolis, Minnesota

ALEXANDER FLAX, President Emeritus, Institute for Defense Analyses, Alexandria, Virginia

ZVI GRILICHES, Professor, Department of Economics, Harvard University, Cambridge, Massachusetts

PHILIP LEDER, John Emory Andrus Professor and Chairman, Harvard Medical School, Cambridge, Massachusetts

GARDNER LINDZEY, President and Director, Center for Advanced Study in the Behavioral Sciences, Palo Alto, California

EDWARD A. MASON, Vice President, Research, Standard Oil Company (Indiana), Naperville, Illinois

DANIEL NATHANS, Professor, Department of Molecular Biology and Genetics, The Johns Hopkins University School of Medicine, Baltimore, Maryland

GILBERT S. OMENN, Dean, School of Public Health and Community Medicine, University of Washington, Seattle

HERBERT A. SIMON, Richard King Mellon University Professor of Computer Science and Psychology, Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania

---

*Term expired December 31, 1984

F. KARL WILLENBROCK, Cecil H. Green Professor of Engineering, School of Engineering and Applied Science, Southern Methodist University, Dallas, Texas

Ex Officio

FRANK PRESS, President, National Academy of Sciences

ROBERT M. WHITE, President, National Academy of Engineering

FREDERICK C. ROBBINS, President, Institute of Medicine

COSEPUP Staff

ALLAN R. HOFFMAN, Executive Director

BARBARA A. CANDLAND, Administrative Assistant

JOANNA MASTANTUONO, Senior Secretary

Preface

In 1982 the Office of Science and Technology Policy (OSTP) asked the National Academy complex[1] to develop a series of briefings aimed at identifying research areas of unusual scientific opportunity within specific fields of science. The first set of briefings covered seven fields, and were presented in both oral and written form to senior officials of OSTP and the National Science Foundation (NSF) in late 1982. The briefings were presented subsequently to senior officials of other interested federal departments and agencies, and published as a collection by the National Academy Press. Responsibility for developing the briefings was given to the Committee on Science, Engineering, and Public Policy (COSEPUP), a joint unit of the National Academies of Sciences and Engineering and the Institute of Medicine.

Thus began the program of annual research briefings that with the publication of this volume will have completed four rounds and covered 28 topics. Topic names and definitions for 1985 are listed below, along with a listing of topics in previous years.

1985[2]

1. *Pain and Pain Management*: Research directed toward enhanced understanding of the basic biological and behavioral phenomena, and their interactions, that underlie pain and analgesia and that lend insight for improved clinical management of pain.

2. *Biotechnology in Agriculture*: Research directed toward increasing the productivity and efficiency of American agriculture using the techniques of modern biotechnology.

3. *Computer Vision and Pattern Recognition*: Mathematical and other research directed toward enhanced capabilities for recognition of objects and geographical locations through comparison of sensor-generated data and stored image information.

4. *Weather Prediction Technologies*: Examination and evaluation of the impact on short- and long-term weather prediction of new

---

[1] Together, the National Academies of Sciences and Engineering and the Institute of Medicine are known as the National Academy complex.

[2] Presentation of briefings on seven topics in 1985 does not imply that other topics not presented are considered to be less important to the progress of U.S. science and technology. No priority setting among fields is implied. The list of 1985 topics reflects the fact that only a limited number of timely topics can be covered in a given year.

technological developments for collection and processing (including mathematical modeling) of meteorological data.

5. *Remote Sensing of the Earth*: Research directed toward enhanced capabilities in all spectral regions for study of the earth from satellite-borne instruments, including observation of land and water surfaces and the atmosphere.

6. *Ceramics and Ceramic Composites*: Research directed toward enhanced understanding of the relationships between microscopic behavior and macroscopic properties of ceramics and ceramic composites, taking into account structure composition and processing history plus environmental conditions to which these materials are exposed. The ultimate goal is to understand these relationships such that ceramic materials can be produced with controlled reproducible properties and their performance under in-service conditions predicted.

7. *Scientific Frontiers and the Superconducting Super Collider*: Examination and evaluation of the scientific opportunities presented by the availability of the proposed superconducting super collider, including opportunities in elementary particle physics and impact on other frontier fields of science.

1984[3]

1. Computer Architecture
2. Information Technology in Precollege Education
3. Chemical and Process Engineering for Biotechnology
4. High-Performance Polymer Composites
5. Biology of Oncogenes
6. Interactions Between Blood and Blood Vessels (Including the Biology of Atherosclerosis)
7. Biology of Parasitism
8. Solar-Terrestrial Plasma Physics
9. Selected Opportunities in Physics

1983[4]

1. Selected Opportunities in Chemistry
2. Cognitive Science and Artificial Intelligence
3. Immunology
4. Solid Earth Sciences
5. Computers in Design and Manufacturing

1982[5]

1. Mathematics
2. Atmospheric Sciences
3. Astronomy and Astrophysics
4. Agricultural Research
5. Neuroscience
6. Materials Science
7. Human Health Effects of Hazardous Chemical Exposures

Topics are generally selected by Dr. George Keyworth, Science Advisor to the President, after consultation with representatives of NSF and COSEPUP. The ceramics topic in 1985 was specifically identified by the NSF as one of timely interest. Each briefing is developed by a panel of experts charged with assessing the status of their field and identifying those research areas within the field likely to return the highest scientific dividends as a result of additional near-term federal investment. Prior to oral delivery of the briefings and their publication they are carefully reviewed by COSEPUP. Financial support for all four rounds of briefings has been provided by NSF.

In its brief history the research briefing activity has become an important channel for communication between the U.S. research community and those responsible for federal funding of research and development. In the words of Dr. Keyworth,[6] "The impacts . . . have been in reinforcing perceptions, in strengthening resolve, and in clarifying the often confusing multiplicity of inputs . . ." converging on the federal agencies. Dr. Keyworth has also pointed out that one of the briefings[7] ". . . led almost directly and quickly to an important new program of Engineering Research Centers in the National Science Foundation." Overall, it is his perception that ". . . the briefings have become increasingly helpful in formulating science policy."

---

[3] Published as *Research Briefings 1984*. Washington, D.C.: National Academy Press, 1984.

[4] Published as *Research Briefings 1983*. Washington, D.C.: National Academy Press, 1983.

[5] Published as *Research Briefings*. Washington, D.C.: National Academy Press, 1983.

Needless to say, the briefings could not have been prepared without the voluntary contributions of the hundreds of scientists and engineers who gave generously of their time. COSEPUP is extremely grateful to those who have served on the briefing panels and helped to develop an important new means of informing government officials.

Leon T. Silver, *Chairman*
Committee on Science, Engineering, and Public Policy

---

[6] *New Pathways in Science and Technology. Collected Research Briefings 1982-84.* New York: Random House, Inc., 1985.

[7] "Computers in Design and Manufacturing," 1983.

Contents

Report of the Research Briefing Panel on Remote Sensing of the Earth 1

Report of the Research Briefing Panel on Pain and Pain Management 19

Report of the Research Briefing Panel on Biotechnology in Agriculture 33

Report of the Research Briefing Panel on Weather Prediction Technologies 47

Report of the Research Briefing Panel on Ceramics and Ceramic Composites 59

Report of the Research Briefing Panel on Scientific Frontiers and the Superconducting Super Collider 73

Report of the Research Briefing Panel on Computer Vision and Pattern Recognition 87

Report of the Research Briefing Panel on Remote Sensing of the Earth

Research Briefing Panel on Remote Sensing of the Earth

D. James Baker (*Chairman*), Joint Oceanographic Institutions, Inc., Washington, D.C.
Francis Bretherton, National Center for Atmospheric Research, Boulder, Colo.
Kevin Burke, Lunar and Planetary Institute, Houston, Tex.
Charles L. Drake, Dartmouth College, Hanover, N.H.
Richard Goody, Harvard University, Cambridge, Mass.
H. William Menard, Scripps Institution of Oceanography, La Jolla, Calif.
Berrien Moore III, University of New Hampshire, Durham
Ronald Prinn, Massachusetts Institute of Technology, Cambridge
John Steele, Woods Hole Oceanographic Institution, Woods Hole, Mass.
Verner E. Suomi, University of Wisconsin, Madison
Wilford F. Weeks, Cold Regions Research and Engineering Laboratory, Hanover, N.H.
Carl Wunsch, Massachusetts Institute of Technology, Cambridge

Staff

Dean P. Kastel, *Staff Director*, Space Science Board
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

Richard Johnson, *Liaison Representative*, Office of Science and Technology Policy

Report of the Research Briefing Panel on Remote Sensing of the Earth

SUMMARY OF RECOMMENDATIONS

Until the flight of satellites, there were no techniques for long-term, global, synoptic measurements of processes in the atmosphere, oceans, and solid earth. Now we are on the verge of establishing a system of remote sensing instruments and earth-based calibration and validation programs that could provide such a data set. With the concurrent development of numerical models that can run on supercomputers, we have the potential of achieving significant advances in understanding the state of the earth, its changes, feedbacks, interactions, and global trends on time scales ranging from days to centuries. To reach this potential, this panel recommends that:

For new missions and new instruments:

High priority be given to the TOPEX mission and the Ocean Color Imager for new starts in FY 1987 and to GRM for a new start in FY 1988, so that simultaneous observations of major processes on the earth will be available by the early 1990s;

Immediate and adequate funding be given to the development of techniques based on satellite remote sensing for the accurate and global measurement of precipitation;

Funds be made available for the development and flight of improved surface imaging spectrometers and tropospheric chemical sensors;

Space be made available on operational missions and Explorer-type earth science missions for testing of new techniques and instruments, and continued and adequate support be provided for the development of an expanded space-based earth observation effort for the 1990s and beyond.

For operational satellite measurements:

The United States as a matter of high priority maintain and continue to upgrade the existing system of civilian operational weather measurements from two polar-orbiting and two geostationary satellites;

Intercalibration of instruments to be used on successive missions that could yield long-term records be an integral part of operational mission planning.

For access of data:

The establishment and maintenance on a continuing basis of adequate, long-term global data bases and correlative observations on specific environmental parameters required to assess both natural variability and human impact on the environment be recognized as a high national priority;

NOAA, NASA, and other appropriate agencies be funded now to adequately develop their data bases and to make them accessible to researchers and other users (This includes archiving the existing satellite data from both U.S. and non-U.S. missions in a form convenient for outside users, and giving proper attention to documenting, storing, reducing, and distributing these data sets);

High priority be given to identifying those data sets of particular importance to long-term changes on the earth, for example, early LANDSAT data, and providing funds to preserve and make available these data to researchers and other users;

Data management be addressed as an integral part of any space mission, and that the views of the interested scientific community be sought in this process to establish requirements for handling, processing, and storing the significant quantities of data expected from future U.S. and non-U.S. missions.

For required complementary measurements and analyses:

The panel emphasizes the importance of a complementary program of earth-based calibration and validation studies, and the critical need for sustaining an overall and coordinated program of space- and earth-based studies together with computer-based modeling. If we are to achieve our goal of understanding the earth system, neither space- nor earth-based measurements can stand alone; the two must be developed simultaneously and tied together with models.

For agency roles and development of a broader community:

The space-based earth science programs of the involved agencies be jointly reviewed and coordinated at the interagency level;

The National Research Council be considered as a channel to provide advice on such a program through the appropriate board and committees;

Basic research funds and facilities be made available as required on a peer-reviewed basis;

Faculty development awards and institutional graduate student support be made available for the development of an adequately trained community.

INTRODUCTION—REMOTE SENSING OF GLOBAL CHANGE

Twenty-five years ago the new meteorological satellites gave atmospheric scientists their first global view of atmospheric processes. Thirteen years ago the first imaging instruments for land processes were flown, and just seven years ago the first ocean satellite provided three months of information on ocean processes and the shape of the geoid. Today satellite remote sensing technology for the earth extends to a wide band of electromagnetic radiation ranging in wavelengths from the microwave to the ultraviolet (see Figure 1).

The emerging global view of the earth and its processes has led us to the verge of dramatic new understanding of the state of planet Earth and its changes on time scales varying from hours and days to months, years, centuries, and millennia. With this new understanding, and with a growing data base on earth processes, we will be able to address the critical questions of predictability and to identify the feedbacks and interactions in the system that give rise to global change. At the same time we will be able to establish the natural variability of the system that is the context for interpretation of developing trends.

Short-term predictability has been treated in other reviews from the Committee on Science, Engineering, and Public Policy: understanding the processes that cause earthquakes by the Research Briefing Panel on Solid Earth Sciences in 1983, and under-

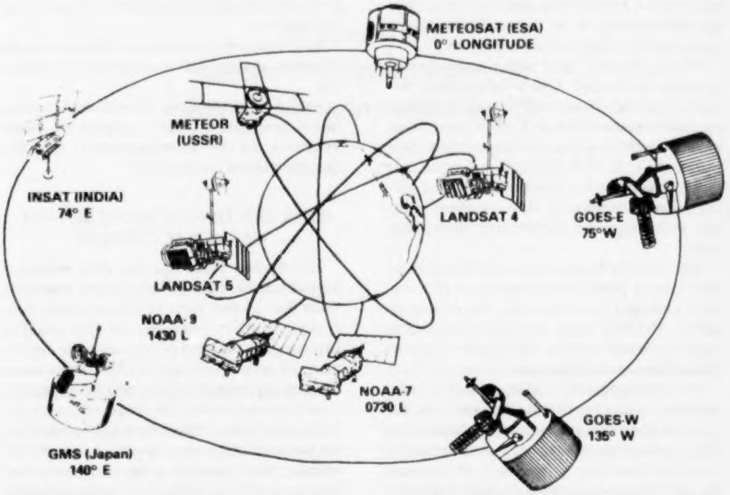


FIGURE 1 Operational Civil Earth Observation Satellites: Polar-orbiting and geostationary satellites monitor weather and surface conditions over the entire globe. SOURCE: NOAA, National Environmental Satellite, Data, and Information Service.

standing the physics of weather prediction on time scales of hours to weeks by the Research Briefing Panel on Weather Prediction Technologies this year. This report identifies special opportunities for advances in understanding global environmental change in time periods ranging from days to centuries. Global change, once it becomes obvious, is likely to be irreversible on human time scales. The ability to diagnose what is happening at the earliest possible stage and to perceive the synergism between seemingly unrelated events may be critical in helping the peoples of the world adjust to changing circumstances and to far-reaching environmental constraints.

The magnitude of global change has only recently become apparent through new remote sensing technology. For example, we now know that the tropical oceans interact with the global atmosphere to cause the El Niño climate anomaly. Ejections of material from volcanoes can lead to global cooling, whereas the increased concentration of gases such as carbon dioxide, methane, and chlorofluoromethanes in the atmosphere can lead to global temperature change. Long-term decreases in upper atmospheric ozone caused by increases in oxides of nitrogen and chlorine can cause increased penetration of ultraviolet radiation, leading to greater incidence of skin cancer, also a global effect. The loss of forests and desertification can be destabilizing, causing climatic anomalies, soil erosion, and consequent loss of agricultural productivity. The cycles of life-supporting elements—carbon, sulfur, phosphorus, nitrogen, and water—are all affected by storage and transfer between the atmosphere, the ocean, the biota, and the solid earth.

To advance our understanding of the causes and effects of global change, we need new observations of the earth. These measurements must be global and synoptic, they must be long-term, and different processes such as atmospheric winds, ocean currents, and biological productivity must be measured simultaneously. We have learned that major advances in earth sciences have come from syntheses of new ideas drawn from such global synoptic observations. The synthesis of plate tectonics from large-scale data is a major step in understanding how the solid earth works; the understanding of the dynamics of large-scale circulation of the atmosphere that comes from global observations has permitted a significant increase in the accuracy of weather prediction. Now we must take the next steps.

Long-term continuity is also crucial. A 20-year time series of the crucial variables would provide a significant improvement in our understanding. Twenty years cover two solar sunspot cycles; it is the period over which we can expect the temperature change due to radiatively active gases to be larger than the natural system noise; it encompasses the eruptions of 5 to 10 volcanoes and the occurrence of 2 to 5 El Niños; and it is the period over which we can expect to see the major effects of deforestation. Finally, we note the need for simultaneity. If we are to make progress in understanding the earth as a system it is essential that we make physical, chemical, and biological observations all at the same time since the physics, chemistry, and biology are all interrelated.

Until the advent of satellites, we had no techniques that could satisfy the needs for long-term, global, synoptic measurements of different processes on the earth. Now we are on the verge of establishing a global system of remote sensing instruments and earth-based calibration and validation programs. Together, these space- and earth-based measurements can provide the necessary data. With the concurrent development of numerical models that can run on supercomputers, we have the potential of achieving significant advances in understanding the state of the earth, its changes, feedbacks, interactions, and global trends on time scales of years to centuries.

NEW MISSIONS AND NEW INSTRUMENTS

It is crucial that data records contain the proper parameters as well as provide long-term continuity. Thus, in addition to the continuance and improvement of the operational weather programs (see below), we need a continuing series of research missions to test new instruments and to measure processes that cannot be defined with data from the limited operational satellite system. The support of a continuing series of research missions is essential to the development of an adequate data set for understanding the earth. The long intervals between launches of research satellites in the earth sciences in the past has not allowed a proper follow-up of earlier developments and has broken the chain of continuous measurements.

With the heritage of the past decade, we now have flight-proven instruments for making global measurements never before undertaken and we can identify areas of development that will make a major change in our understanding of the state of the earth. It is essential that there be a long-term national commitment to satellite missions that will carry the instruments that are now ready, and we must now begin to develop the necessary new instruments and techniques.

New missions identified as high priority are the Ocean Topography Experiment (TOPEX) and the Geopotential Research Mission (GRM). An important new instrument ready for flight is the Ocean Color Imager. Areas of development include spectral imaging instruments for land biology and geology and for tropospheric gases, particularly carbon monoxide, and new techniques for measurement of precipitation. In addition, the panel sees important opportunities for Explorer-type missions for earth science; two of particular interest are the measurements of the large-scale magnetic field and the test of tropical precipitation and tropospheric carbon monoxide measurement techniques.

THE OCEAN TOPOGRAPHY EXPERIMENT (TOPEX)

The short-lived SEASAT mission showed the oceanographic community that a technological revolution in ocean observation was finally possible. For more than 100 years, oceanography has had to rely upon ships and isolated, short-duration instruments to gain information about a global, turbulent fluid. Until the advent of the sensors on SEASAT, the density of observations possible over and within the ocean necessarily failed by several orders of magnitude to meet even the most rudimentary of the known sampling requirements of a turbulent fluid, one where all scales of motion were known to be important.

SEASAT showed that modern microwave electronics could provide global, continuous measurements of surface atmospheric wind stress, the most important field of forcing of the ocean, and of the dynamical response of the ocean through the movements of its upper surface. Plans are now in place for the Navy Remote Ocean Sensing System (NROSS) and the European Space Agency's ERS-1 to provide the wind stress measurements beginning in 1990; the precision ocean topography measurements needed for ocean currents will be provided by the proposed NASA mission TOPEX.

The U.S. and world oceanographic communities have spent the past five years constructing an orderly approach to getting the new observational technologies into place. Much of the planning has taken place in the context of the World Climate Research Program, which includes the Tropical Ocean-Global Atmosphere Program (TOGA) and the World Ocean Circulation Experiment (WOCE). WOCE planning, which includes global satellite and *in situ* programs of ocean measurements, has led to an orderly sequence for the proposed launch of the requisite spacecraft (NROSS, TOPEX [in its most desirable form, as the U.S./French mission TOPEX/POSEIDON], GRM [see below], and ERS-1). The flight of these spacecraft, phased approximately over the period 1990 to 1996, would be accompanied by the international application of all the modern physical and chemical oceanographic techniques on a global scale.

Figure 2 shows the dramatic improvement in the estimate of the meridional heat flux in the ocean anticipated with the application of the combination of TOPEX altimetry and WOCE tropical carbon-14 tracer data.

THE GEOPOTENTIAL RESEARCH MISSION (GRM)

The Geopotential Research Mission is designed to measure variations in the gravity and magnetic field over the entire earth to a resolution of 2 milligals and 2 nanotesla for 100-km half wavelengths. Because it will provide a much improved geoid, to which sea surface heights are referred, the GRM is needed to allow full realization of the results of the contemporary TOPEX mission. GRM is proposed to fly during the NROSS/TOPEX flight period.

Gravity measurements from GRM are expected to reveal, for the first time, the pattern of underlying mantle circulation. This information will show how the smaller scales of mantle convection interact with the large-scale plate structure of the lithosphere. Tantalizing patterns of mantle circulation beneath the oceans have been discerned from SEASAT altimetry, and seismic tomography is beginning to reveal the deeper structure of the mantle. As the global digital seismic network (GSN) becomes operational, the data

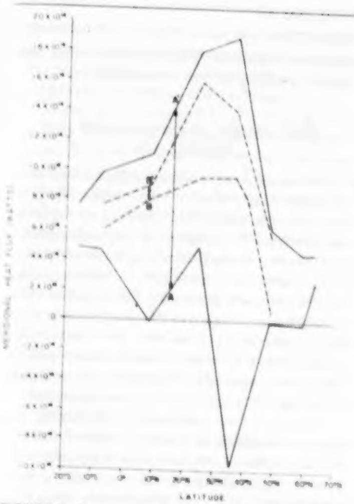


FIGURE 2 Expected improvement of estimate of ocean heat flux with new data sets: A-A' is the estimated uncertainty for the annual average meridional heat flux in the North Atlantic using existing data; B-B' is the uncertainty expected with the application of satellite altimetry and *in situ* carbon-14 tracer data. SOURCE: C. Wunsch, Massachusetts Institute of Technology.

sets from GRM, SEASAT, and GSN can be combined to reveal for the first time the relation between plate motion and the underlying mantle circulation.

Gravity fields measured from GRM will uniquely illuminate the mass distribution under the Himalayas, Tibet, and neighboring areas of Asia where the collision of India with Eurasia is currently being accommodated by the elevation of the world's highest mountains and by the extrusion of China toward the Pacific Ocean. Collisions of this kind have been a major element in the evolution of the continents over the last four billion years; acquisition of the GRM gravity field is essential to full description and understanding of this complicated process. A complementary data set will be acquired by GRM over the high Andean chain where calculated mass distributions can be used to establish how active subduction and volcanism operate.

The 160-km altitude orbit, essential for the GRM gravity measurements, also provides an unprecedented opportunity to fly a magnetometer in very low earth orbit. This instrument will not only measure the large-scale magnetic field, but will also establish the distribution of anomalies in the field at the 100-km scale that are attributable to ferromagnetic minerals in the continental crust. Over much of the continents this information is otherwise unattainable; thus acquisition of the GRM magnetic field represents an essential element in establishing earth history. A magnetic field explorer mission, at higher altitudes, will measure the large-scale magnetic field for studies of its generation and changes.

THE OCEAN COLOR IMAGER

A major concern for earth science in the next decade and beyond is an understanding of biogeochemical cycles; at decadal time scales the cycles play a major part in critical processes such as the carbon, nitrogen, and sulphur exchanges. The oceans' role in these cycles is known qualitatively and operates in part through vertical exchanges: there is strong and rapid coupling between surface and deep water through particle fluxes that carry organic matter. Large-scale connections were recently seen in the changes in atmospheric carbon dioxide occurring at the time of the 1982–1983 El Niño, which produced massive changes in physical, chemical, and biological properties in the Pacific.

Investigations of these vertical fluxes are under way through studies of chemical distributions in the ocean; by direct measurements of particle transports and the rates of transformation between particulate and soluble phases; and through modeling of the combined physical, chemical, and biological dynamics. These studies depend on a knowledge of the primary production of organic matter in the upper layer of the ocean. We have methods to determine these rates by shipboard experiments at discrete locations, but we require a global perspective of this process and its variability in space and time.

Satellite measurements of ocean color can provide this synoptic view. The Coastal Zone Color Scanner (CZCS) was intended for regional (and fisheries) studies, but has also proved capable of use for open ocean study. Algorithms have been developed to relate the remotely sensed data to chlorophyll, the primary pigment for photosynthesis. In turn, these estimates can be correlated with measurements of primary production. The proposed Ocean Color Imager (OCI), with more spectral channels, will provide more accurate measurements of chlorophyll. The new data will yield a better understanding of the essential links between physical changes in the ocean, biological production (including fisheries), and chemical fluxes to deep water. The integration of these processes using satellite data is essential for furthering our understanding of general biogeochemical cycles. It will be possible to test hypotheses about the role of these cycles, not only in their coupling with ocean dynamics but also in controlling the overall planetary system, including feedback to atmospheric and terrestrial components.

These concepts and the necessary technologies are now being formulated into a plan for a Global Ocean Flux Study, which can provide the links between physical oceanography programs and the corresponding long-term atmospheric and terrestrial programs. The studies on basic production are relevant to a further initiative on recruitment processes, and the OCI is an essential and central element in these overall plans.

MEASUREMENT OF PRECIPITATION AND EVAPORATION

Accurate measurements of the rates of precipitation, evaporation, and evapotranspiration over the global land and ocean surfaces are essential for understanding the coupled earth-land-ocean-atmosphere system. Evaporation is closely tied to temperature, wind stress, and humidity. Strong efforts are being made to improve these measurements from remote instruments. Unfortunately, precipitation is highly variable and poorly measured especially over the ocean; we lack a reliable technique for this measurement.

There are two problems in obtaining the global coverage of rainfall desired for climatic studies: sampling and measurement. The traditional source of rain data, a rain gauge network, is unavailable over the oceans. Because of the convective nature of the rainfall, island stations are not representative of the surrounding ocean, and ship station rainfall observations are sparse as well as notoriously unrepresentative and unreliable.

During the Atlantic Tropical Experiment of the Global Atmospheric Research Program it was shown that satellite microwave data were highly correlated with calibrated ship radar when both were observing the same area. Unfortunately the satellite systems were turned on only every other orbit and had a swath width less than the interorbit gap. Thus, while microwave multifrequency radiometers on polar-orbiting satellites appear to meet the measurement accuracy criteria, they fail on adequate sampling.

Geostationary satellites sample much of the earth continuously and the tropical belt completely. They provide information on where it is raining, but not on how much. Several rain estimation schemes have been devised; for example, there is good evidence that the cloud height as measured by geostationary satellites is closely related to rain rate for convective rain. Unfortunately, the errors in this technique are too large.

There are a number of promising techniques that need support for development now. Such techniques follow two schemes, scattering or absorption. The scattering scheme, useful at the higher microwave frequencies, is largely empirical. On the other hand, the theory of microwave attenuation due to rainfall is in much better order. Here the key uncertainties remaining are the thickness of the rain layer and the presence or absence of the ice phase. A satellite equipped with a suitable radar can remove this uncertainty and greatly improve the accuracy of microwave observations. Adequate sampling from a polar orbiting satellite would still be a problem.

In summary, the requirements for accurate rainfall measurement are demanding, but we believe that satellite platforms now in orbit could be equipped to test and improve our measurement capability. The development and flight of precipitation measurement techniques on operational satellites and on Explorer-type missions is of high priority for any earth sciences program.

Measurements With Improved Imaging Spectrometers

The most broadly useful instrument for geology and biology among the new capabilities envisioned for Earth observation in the 1990s is a moderate-resolution imaging spectrometer (MODIS). The spatial resolution would be similar to that presently available (about 1 km), but the focus would be on greatly increasing the spectral information. We now know that important information about the earth and its biology can be found by observing precisely and simultaneously in many parts of the visible and infrared spectrum. For example, vegetative stress, including chemical imbalances induced from acid rain, can be detected as can the subtle composition of minerals through the spectra of hydrous compounds.

The MODIS instrument would also offer a major enhancement to ocean color measurements begun by the Coastal Zone Color Scanner and to be extended by the Ocean Color Imager, and it would greatly expand the capabilities offered by the current operational radiometers. The new technique is now available and ready for development for use in space. As part of this development there is a need for experimental use of the spectral techniques over test areas.

These techniques promise a new dimension for remote sensing from space for geologic structure and geomorphology that has been successful since the time of the first LANDSAT mission (1972). Continuing improvements in technique (e.g., stereo images, nonnadir look angles, higher spectral resolution, imaging at a variety of times of day) are likely to lead to increasing demand for the use of these tools and also of related spaceborne photography (e.g., large-format camera products).

Tropospheric Chemical Sensors

We now know that the chemistry of the troposphere is controlled to a considerable extent by biological and human processes and hence is not static but changing. Satellites have played important roles in the study of upper atmosphere chemistry and will continue this role with the launch of the Upper Atmosphere Research Satellite (UARS) in 1988. In contrast, satellites have not yet played a major role in tropospheric chemistry because of difficulties in remote sensing of many tropospheric species. However, carbon monoxide, which is one of the most important species controlling the oxidation state of the global atmosphere, is detectable from space. We urgently need to know the global distribution and trend of carbon monoxide, but this gas is so variable in space and time that satellites appear to provide the only practical way of meeting this need. An exploratory sensor has been flown on the space shuttle, which shows the feasibility of CO measurements from space. A more sophisticated sensor capable of measuring the vertical profile of CO over the globe and flown on a polar orbiter would provide data of crucial importance to our understanding of global tropospheric chemistry.

The panel notes the importance both of the new research missions and of the development of new techniques and instruments to build a base for an expanded space-based earth observation effort for the 1990s and beyond. Specifically, the panel recommends that:

- *High priority be given to the TOPEX mission and the Ocean Color Imager for new starts in FY 1987 and to GRM for a new start in FY 1988, so that simultaneous observations of major processes on the earth will be available by the early 1990s.*
- *Immediate and adequate funding be given to the development of techniques based on satellite remote sensing for the accurate and global measurement of precipitation.*
- *Funds be made available for the development and flight of improved surface imaging spectrometers and tropospheric chemical sensors.*
- *Space be made available on operational missions and Explorer-type earth science missions for test of new techniques and instruments, and continued and adequate support be provided for the development of an expanded space-based earth observation effort for the 1990s and beyond.*

OPERATIONAL WEATHER AND ICE MEASUREMENTS—THE RESEARCH BASE

The U.S. civil operational earth-observing satellite systems, together with their earth-based calibration and validation programs, provide a global data base for meteorology, oceanography, solid-earth geophysics, and solar-terrestrial science. From these sources, federal agencies provide operational services critical to the protection of life and property, the national economy, energy development and distribution, and global food supplies.

At the same time, this data base is fundamental to research on the state of the earth and its predictability. Thus, it is essential that the operational satellite measurement system be continued and improved to provide homogeneous, accurate, and timely data. The panel notes that its recommendations, as well as those from the Panel on Weather Prediction Technologies, are based on the assumption that the operational systems now in place will continue.

Today, two U.S. polar-orbiting weather satellites and two U.S. geostationary satellites monitor weather and surface conditions all over the globe. Two U.S. earth resources satellites (LANDSATs) are also in polar orbit. These U.S. systems are complemented by a number of foreign Earth-observing satellites, and more foreign satellites will be launched in the latter half of the 1980s.

The polar-orbiting weather satellites provide operational coverage of the entire earth four times per day, measuring temperature and humidity in the Earth's atmosphere, surface temperature, cloud cover, water-ice boundaries, and proton and electron flux near the earth. They receive, process, and distribute data from stations distributed globally. Derived polar satellite products include information on ice, vegetation, snow, and tropical cyclones.

Improvements needed for future polar-orbiting satellites include additional channels in the sounding units and added capability in the radiometers for better atmospheric moisture measurements, better estimation of worldwide vegetation levels, and operational measurement of solar and terrestrial radiation levels.

The geostationary satellites provide synoptic hemispheric coverage that extends well into the southern hemisphere. Their sounding/imaging instruments make day and night observations of weather in the coverage area and relay data from surface collection points to processing centers. The geostationary sounders provide a complementary measurement to the polar-orbiters with their synoptic vantage point, and also monitor the magnetic field of the earth, the energetic particle flux, and gamma ray emissions from the sun.

Improvements in the next generation of geostationary satellites include replacement of the current sounding/imaging instrument with separate instruments. A stabilized spacecraft would increase the available scanning time, and detector arrays rather than single-point detectors could scan swaths of the earth and so reduce the time for complete area coverage while retaining the same size field of view and multichannel spectral data input.

Lack of conventional observational stations in the polar regions still represents a major data gap. The synoptic observations required are only possible with satellite remote sensing. Present satellite systems have already contributed significantly to improved observations even though sensors in the visible and infrared are limited by clouds and darkness. The most useful remote sensing system for sea ice studies currently deployed is a passive microwave imager which provides all-weather, low-resolution imagery. However, an adequate validation of the algorithm, or transfer function, used to calculate ice conditions from this data is still lacking.

The next step in ice-observing capabilities will come with the launching of an improved passive microwave system on a U.S. Defense Meteorological Satellite in 1986. Then we look to the flight of synthetic aperture radar systems on European, Canadian, and Japanese satellites during 1989 to 1995. For the United States to receive maximum benefit from the radar systems, it will be necessary to establish ground truth observations during shuttle test missions in 1987 and 1988 to validate the interpretation of this imagery over ice-covered oceans. In addition, the schedule for establishment of receiving stations for the non-U.S. satellites must be maintained. The high resolution of synthetic aperture radar (tens of meters) makes it valuable also for open ocean and land surface studies.

The need for all these operational measurements has been firmly emphasized internationally. At the seventh meeting of the Economic Summit of Industrialized Nations the heads of state established a Working Group on Technology, Growth, and Employment; remote sensing from space is the topic of a panel chaired by the United States. The summit panel has already emphasized the importance of long-term continuity of operational meteorological satellite data.

In view of the fundamental importance of the data that is collected by the operational weather satellites to the research community, the panel recommends:

- *That the United States as a matter of high priority maintain and continue to upgrade the existing system of civilian operational weather measurements from two polar-orbiting and two geostationary satellites.*

In addition, to ensure that long-term data records are as homogeneous as possible, the panel recommends that:

- *Intercalibration of instruments to be used on successive missions that could yield long-term records be an integral part of operational mission planning.*

ACCESSING THE DATA

Over the next 10 to 15 years we expect to see a major increase in data volume from the operational remote sensing systems and the new research satellite programs, coming largely from requirements for higher spatial and spectral resolution, repetitive measurements, long-term data sets, and global processes. At the same time, we find now that access to existing remotely sensed data is not adequate for research purposes. Thus a major high-leverage investment could be made now in the improvement of the data management, archiving, and distribution of operational satellite data for research users.

The effective management of large data bases is a general problem in space science, and has been addressed in a number of reports. In essence, it has proved almost impossible for the researcher to "drink from the firehose" of satellite data streams. It will be far more difficult to digest data from a number of different sources. In the panel's view, the scientific exploitation of Earth observations are likely to be limited as much by the inability to manipulate and distribute efficiently large and diverse data sets as by sensor technology. The challenge lies in information development and distribution; a challenge that can be met with new high-technology computing and communication techniques if adequate funds are made available. The several planning efforts now under way for high-speed data networks to link supercomputers, data centers, and university research laboratories need continued support.

The panel's recommendations for data management are in two parts: one on current data, and one on data that will come from satellites that will fly in the next decade. In terms of current data, the panel recommends that:

- *The establishment and maintenance on a continuing basis of adequate, long-term global data bases and correlative observations on specific environmental parameters that are required to assess both natural variability and human impact on the environment be recognized as a high national priority;*
- *NOAA, NASA, and other appropriate agencies be funded now to develop their data bases adequately and to make them accessible to researchers and other users (This includes archiving the existing satellite data [from both U.S. and non-U.S. missions] in a form convenient for outside users, and giving proper attention to documenting, storing, reducing, and distributing these data sets);*
- *High priority be given to identifying those data sets of particular importance to long-term changes on the earth, for example, early LANDSAT data, and providing funds to preserve and make available these data to researchers and other users.*

Looking ahead, the panel supports the recommendations of the Space Science Board's Committees on Earth Sciences and on Data Management and Computation that:

- *Data management be addressed as an integral part of any space mission, and that the views of the interested scientific community be sought in this process to establish requirements for handling, processing, and storing the significant quantities of data expected from future U.S. and non-U.S. missions.*

Specific developments that need funding in the next two years include high-speed parallel processors and other large computational machines that will be required for the large data sets to come, and the development of communications networks linking data centers, supercomputers, and university research laboratories.

REQUIRED COMPLEMENTARY MEASUREMENTS AND ANALYSES

The measurements obtained from satellites must be complemented by measurements in the field to reliably calibrate, validate, and interpret the satellite data. In addition, many of the important scientific problems that can be addressed using satellite measurements also require measurements taken on the surface of the earth or from aircraft. It is essential that the program for remote sensing of the earth from space include such complementary field measurements.

In the atmosphere, measurements of basic meteorological variables (temperature, pressure, humidity) from balloons have long been recognized as essential to the accurate interpretation of satellite data. In addition, while satellite data are useful for the study of storms, ground-based and airborne measurements provide data on variables such as wind velocities at levels of accuracy that cannot be attained from satellites.

Also in the atmosphere, many chemically important trace species are not measurable from space with current techniques. However, the concentrations of these species can be obtained through current ground-based or airborne techniques. Such information is essential to understanding the chemistry of species such as stratospheric ozone and tropospheric carbon monoxide, which are detectable from satellites. Trends of very long-lived gases are usually best measured at the surface. In addition, surface vegetation and soil and oceanic microorganisms are dominant sources of many chemically and radiatively important trace gases, and field measurements are the only way to detect and quantify these emissions.

In the ocean, data from acoustic sounders, *in situ* measurements of the vertical profiles of temperature and salinity, and velocity and other data from buoys are critical complements to satellite data on sea surface topography and temperatures for the study of oceanic heat and momentum fluxes. In addition, *in situ* measurements of primary production, nutrient concentrations, and downward fluxes of organic material are essential for quantitative interpretation of satellite measurements of ocean color and provide the necessary data to examine the very significant role of the oceans in the global carbon cycle.

For the land surface, intensive field studies of major vegetative groupings (biomes) will be necessary to define the water, carbon, nitrogen, phosphorus, and sulfur cycles within these biomes. Such information is essential for interpreting remotely sensed data on areal extent and vegetative density in terms of the global hydrologic and biogeochemical cycles.

Finally, the satellite data on geopotential fields will prove most useful when they are combined with other, land-based geophysical data to decipher deep earth structure and global tectonic processes.

In situ observations for calibration and environmental characterization are equally important in the polar regions. However, because of the hostile nature of the environment as well as the isolation, continued manned measurements are both difficult to arrange and expensive. One answer to this problem is the deployment of extensive data buoy arrays on the drifting ice. In addition to improved measurements of standard oceanographic and atmospheric parameters, observations should also be possible of internal ice stress, ambient underwater sound, water mass movements, and water column properties. Such an expansion in data collection abilities would clearly have to be accompanied by an associated expansion in satellite data transmission capabilities.

The panel emphasizes the importance of a complementary program of earth-based calibration and validation studies and the critical need for sustaining an overall and coordinated program of space- and earth-based studies together with computer-based modeling. If we are to achieve our goal of understanding the earth system, neither space- nor earth-based measurements can stand alone; the two must be developed simultaneously and tied together with models.

AGENCY ROLES AND THE BROADER COMMUNITY

It can be said that remote sensing is no more than a tool; yet it is a revolutionary tool for the earth sciences, offering global, continuous, and long-term data gathered with the same instruments over all the oceans and all the continents. The panel emphasizes that this tool has yet to be fully exploited and notes that this will take decades of future effort by leading researchers. The now-traditional *in situ* techniques required a similar sustained effort from previous generations.

Ease of transition from technological innovation to use, both operationally and in support of fundamental research, requires that the process be treated as a whole, a requirement not well suited to the division of responsibility existing among the federal agencies. Sensor development is usually NASA's responsibility, long-term data collection that of NOAA and DOD, while fundamental research into global systems lies closest to the mission of NSF.

AGENCY ROLES

The lack of a unified approach to remote sensing among the federal agencies is a long-standing issue. To resolve this problem the panel proposes that review and oversight at the interagency level be provided for the programs and budgets for space-based earth measurements of all the agencies involved. This review and oversight activity, following precedents such as the National Climate Program and the National Acid Precipitation Assessment Program, should continue for a decade, have a wide scope, and involve a broader scientific community than that which exists within the government.

There are many concerns about the proper use of remote sensing that could profit from interagency oversight: coordination of agency activities and budgets; use of operational satellites for research; impacts of classification; coordination with surface and *in situ* measurements; data distribution and use; involvement of the nongovernmental science community; and international collaboration.

Our preference is to allow the characteristics of the oversight to evolve from continuous discussions between agencies and the science community rather than to attempt rigid definitions of agency roles. But all agencies, in particular, NSF, must be involved in the oversight process, in recognition of the growing importance of remote sensing to many earth science activities.

The research community also must be strongly represented, even perhaps in a majority; the National Research Council could be an effective channel. We recommend that such a channel be developed.

THE BROADER SCIENTIFIC COMMUNITY

Remote sensing cannot achieve its potential as a research tool until it has been thoroughly used to explore at the frontiers of earth science. Remotely sensed parameters do not arrive at the desk of a research scientist as if they were tables of lengths, masses, or temperatures. Instead, an established scientist is likely to be faced with a long period of self-education during which he or she may achieve few tangible results. Many have faced this choice and decided that it involves too great a risk for their career prospects, even though they are convinced that the use of space data is of great benefit to the subject. Federal agencies can ease this situation.

Specific examples of such support include faculty development awards to allow senior faculty and staff to carry out special study and to go on leave, and to provide promising young faculty with the opportunity to develop facility in the use of the new techniques and data sets. Institutional grants for training of graduate students would also enable the United States to provide the personnel needed to capitalize on this new technology.

In terms of basic research support, we propose that NASA, NOAA, DOD, and NSF work together to set aside funds to support research in areas with a large remote-sensing component and that they signal their intention by periodic announcements of opportunity. Two matters are of primary concern: that funded proposals be of high quality measured against those of the parent community, not merely a subcommunity addressing remote sensing; and that the continuity of funding be ensured.

RESTRICTIONS ON DATA USE

From time to time, especially as data about the earth become more and more accurate and detailed, there are pressures from both national security interests and from the industrial community to restrict various classes of information. The panel notes the need for the United States to map and make available data on the new resource of the Exclusive Economic Zone, and to provide satellite techniques like the Global Positioning System, which has enormous promise for allowing measurements of location and rates of crustal movements that we previously have only been able to infer. At the same time, the panel recognizes the need to restrict data for national security reasons. We urge that such restriction be imposed only after careful thought as to the exact mode of restriction, and that scientists who are familiar with the research needs of the earth sciences community be involved in such decisions. It would be a great national loss if the U.S. research community were to surrender its world leadership for arbitrary reasons of classification.

In summary, the panel recommends that:

- *The space-based earth science programs of the involved agencies be jointly reviewed and coordinated at the interagency level,*
- *The National Research Council be considered as a channel to provide advice on such a program through the appropriate board and committees,*
- *Basic research funds and facilities be made available as required on a peer-reviewed basis,*
- *Faculty development awards and institutional graduate student support be made available for the development of an adequately trained community.*

CONCLUSIONS

Knowledge of the earth is fundamental to humanity's survival. Satellite-borne instrumentation has provided scientific data that, when used in conjunction with *in situ* measurements and theoretical models, has given us a glimpse of the knowledge that could be gained with full application of these techniques. Thus we are now ready to measure, on a long-term basis, the essential processes that govern the state of the earth and its changes over periods of years to centuries.

The first step in this process is the flight of research missions with proven technology to measure those processes not now covered by operational missions, such as ocean physical and biological processes and the global gravity and magnetic field. Together with these research missions we need to build a base for an expanded earth observation effort for the 1990s and beyond. In particular, we need to develop an accurate global technique for measuring rainfall, to provide an imaging spectrometer that will reveal new features of geology and biology, and to develop tropospheric chemical sensors. In addition, we need to continue the existing operational weather satellites, to make the existing data available in a useful form to researchers, and to save our treasure of older data that is now being lost. With all this in place, we will have established the beginning of a global earth measuring system.

The space missions and new instruments are not enough in themselves, however. Adequate data systems must be in place that use state-of-the-art technology for archiving, storage, and accessibility. Major programs of *in situ* measurements including both calibration and validation and process studies will be required; these must be funded at the same time. The data systems and the theoretical models that use the data both will call for the most powerful supercomputers; thus the earth science community adds its voice to that need.

Finally, the panel notes the importance of interagency cooperation in earth science. Collection of data, calibration and validation, archiving and dissemination, and support of basic research cut across many agencies in the federal government. Moreover, the industrial and national defense communities have a real and long-term interest in much of this information. As a consequence, if the United States is to keep its lead in this area of major importance close interagency cooperation is required. The panel recommends that the appropriate oversight be provided at the interagency level.

REFERENCES

The panel has based its recommendations and identification of priorities on relevant reports by committees of the NAS/NRC, especially the reports of the Committee on Earth Sciences of the Space Science Board, *A Strategy for Earth Science from Space in the 1980s and 1990s, Part I: Solid Earth and Oceans*, and *Part II: Atmosphere and Interactions with the Solid Earth, Oceans, and Biota*. The panel has also used the reports of the Committee on Data Management and Computation of the Space Science Board, *Data Management and Computation: Volume 1: Issues and Recommendations*, and *Volume 2: Space Science Data Management Units in the 1980s and 1990s*. These strategy reports are part of a long-term effort by the Space Science Board to set scientific goals and priorities intended to maximize the scientific return on the nation's investment in space science.

In addition, the panel has drawn on the ongoing deliberations of the NASA Earth Systems Sciences Committee, the NASA Earth Observing System's Science Working Group, and the Ad Hoc Group on Remote Sensing for Global Change of the Committee on Space Research of the International Council of Scientific Unions for help in identifying the highest-priority recommendations.

Report of the Research Briefing Panel on Pain and Pain Management

Research Briefing Panel on Pain and Pain Management

Edward R. Perl (*Chairman*), University of North Carolina, Chapel Hill
Benjamin G. Covino, Harvard Medical School and Brigham and Women's Hospital, Boston, Mass.
Howard L. Fields, University of California Medical School, San Francisco
Kathleen M. Foley, Memorial Sloan-Kettering Cancer Center, New York, N.Y.
Wilbert E. Fordyce, University of Washington and University Hospital, Seattle
John A. Jane, University of Virginia School of Medicine, Charlottesville
Thomas Jessell, Harvard Medical School, Boston, Mass.
Donlin M. Long, The Johns Hopkins Hospital, Baltimore, Md.
David U'Prichard, Nova Pharmaceutical Corporation, Baltimore, Md.
Charles J. Vierck, Jr., University of Florida College of Medicine, Gainesville

Staff

Elizabeth Bullitt, *Rapporteur*, University of North Carolina, Chapel Hill
Barbara Filner, *Director*, Division of Health Sciences Policy, Institute of Medicine
Naomi Hudson, *Administrative Secretary*
Allan R. Hoffman, *Executive Director*, Committee on Science, Technology, and Public Policy

We thank The Neurosciences Institute, New York, N.Y., for making its facilities available for the panel meeting.

Report of the Research Briefing Panel on Pain and Pain Management

INTRODUCTION

What is pain? This question has troubled philosophers and scientists since before Aristotle's time. Pain has attributes of a sensation, yet the regularity with which it makes us suffer or feel uncomfortable distinguishes it from touch and other sensations of the body. Is it then an emotion? This uncertainty about the exact nature of pain is reflected in our language. We use the term "pain" for the feeling produced when we are physically injured and also for our emotional reaction when we suffer as a consequence of unkind words or loss of a loved one. Confusion about meaning has contributed to difficulties in research and treatment.

Pain after bodily injury is a physiological phenomenon, the perceptual part of the normal response to a variety of stimuli that are physically threatening to the integrity of tissues, i.e. noxious. Because of the aversive feature of discomfort or suffering, pain as a sensory experience has evident protective value.

Pain develops another dimension when it persists or occurs repeatedly. The discomfort and suffering associated with it disturb the quality of life and can produce complex and profound alterations in behavior. As a consequence, pain is a major health problem. It is the single most common complaint causing people to seek health care. People worldwide take enormous quantities of drugs because of pain. Low back pain alone accounts for well over 10 million physician visits annually in the United States. All of this makes pain a formidable social and economic problem.

THE TWO SIDES OF RESEARCH ON PAIN

Research on pain and its mechanisms has two somewhat different but complementary aims. One seeks to understand the basic nervous mechanisms underlying pain, including their link to behavior and other neurally mediated phenomena. The other investigates the control and management of pain in pathological and clinical situations.

Research on the neural mechanisms of pain has contributed new ideas about how the brain works. For example, the traditional concept has been that the operation of complex nervous systems could be understood largely in terms of the connections or "wiring" of the constituent cells activated by the circulation of brief electrochemical events. Through study of pain mechanisms, it is now known that nerve cells employ chemicals to mediate signals other than those traditionally associated with such function. Study of the newly discovered agents has shown that long-term, chemically mediated interactions between nerve cells are part of the process by which neural actions are produced. Complex control mechanisms by which sensory information is modulated and channeled have also been brought to light.

Improvements in the management of pain have also come about as a consequence of better understanding of its neural basis. Other changes in the therapeutic approaches have resulted from enlightened appreciation of the link between pain and complex behavioral factors. New strategies for pain management include, for example, deliberate stimulation of nervous pathways that are part of the control system for pain mechanisms. Means for controlling pain by drugs while reducing risk of addiction and new categories of drugs useful in such control have been proposed based upon the emerging understanding of the chemical messenger systems in the brain. Also, behavioral testing and behavioral management of patients with chronic pain has become a recognized part of clinical strategy.

Nevertheless, these developments in understanding pain mechanisms and pain control represent only a beginning. The work to date is probably most significant in that it has opened the door to future developments rather than provided definitive answers to the basic scientific questions or to clinical problems.

THE SHAPE OF NEW KNOWLEDGE ON THE NEURAL BASIS OF PAIN

Discoveries about the neural substrates for pain have reshaped our thinking about it as a sensation and altered the direction research in the field will take. Special sense organs for pain have been found and they have been shown to feed more or less dedicated neural pathways. These dedicated networks are paralleled by less specific neural systems whose activity combines input from the pain receptors with other messages from the body's tissues. The neural networks activated by pain-producing stimulation have been found to use a diverse set of chemical messengers to transmit activity and modulate the sensitivity of the nerve cells making up the network. These chemical mediators produce their effects by binding to selective receptor molecules on the surface of target neurons. Moreover, researchers have uncovered a modulatory system whereby the brain alters the flow of activity through the neural networks transmitting pain-related signals.

The practical importance of these discoveries is substantial. Knowledge about specific neural networks for pain has provided a focus for work on pain control. As previously stated, alternate clinical strategies based upon deliberate manipulation of the neural control systems for the networks related to pain are being explored, and information on the neural chemical mediators and the receptor molecules on which they act provides new opportunities for the development of drugs to better control pain while minimizing undesirable effects such as habituation and addiction. This report comments on the general nature of pain as a sensation, the peripheral and central nervous system mechanisms for pain, related behavioral factors, and clinical considerations.

THE SENSE ORGANS FOR PAIN AND CHEMICAL INTERMEDIARIES

There are certain common features in neural arrangements for the various sensations in human beings and other mammals. The starting point is a feature or modification of the environment—a stimulus, e.g., light, sound, tissue swelling. The stimulus must be transduced into nervous system messages. The sense organ where this transduction takes place is more or less specialized to respond to a particular subset of physical and chemical changes.

Sense organs usually are located some distance from the central nervous system. Like most other nerve cells, those leading from the sense organs to the central nervous system have an elongated extension, a fiber, over which the nerve cell activity must be transmitted. In human beings the distance the activity must travel along the fiber can be from fractions of a millimeter to more than a meter. The information represented by the activity also must be conveyed from one nerve cell to the next.

Nerve cells transmit activity across themselves to distant parts by a brief electrochemical event, the nerve impulse. The repetition interval between successive nerve impulses represents the code for information about the activity.

Activity is transmitted between cells at special junctions called synapses, where the membranes of two cells come into close apposition. Commonly, a chemical mediator released by the presynaptic cell (sending cell) diffuses across the tiny gap (about 200 angstroms) to affect the postsynaptic cell (receiving cell) by combining with a receptor specialized to bind a particular chemical configuration. The impulse code and transfer at synapses is how activity generated by the sense organ is conveyed across the various neural elements of the network.

For all of the perceptual reactions accepted as sensations, activity initiated by the sense organ(s) passes through a series of more or less dedicated groups of neurons at a succession of stations in the brain to the level of the brain's cerebral cortex. Ultimately, the activity of these specialized neural networks is translated into conscious experience with its concomitants of discrimination, feature detection, and quantitative recognition.

Until relatively recently, direct evidence for this kind of organization did not exist for pain. The nature of the sense organ, i.e., the transducer, was not understood, although it was known that conduction by a subpopulation of the nerve fibers in peripheral nerves was necessary for pain to be provoked by the usual stimuli. It was also understood that a particular bundle of nerve fibers in the spinal cord, a pathway called the spinothalamic tract, was essential for transmitting information to the brain about pain-causing stimuli. However, knowledge about the origin of this pathway, the signaling characteristics of the neurons that formed it, and the nerve net with which it articulated at higher centers in the brain was sketchy and in dispute.

Sense organs for pain Whether or not there were specific sense organs for pain was a major controversy for over a century, a controversy that dominated theories about the basis of pain. During the past 15 years, a set of sense organs whose activity produces pain has been documented. Earlier difficulties in establishing their existence was largely due to technical problems related to the small size of the neurons involved. In general, pain receptors have thin fibers in peripheral nerves. This special set of sense organs in the body has substantially higher thresholds (requires more intense stimulation) for activation than other sense organs in the same tissues responsive to mechanical, temperature, or other nonpainful stimuli.

Because of their high thresholds, the weakest stimuli exciting the sense organs for pain approach or exceed the intensity that can damage tissue. Thus, sense organs for pain can be considered noxious stimulation receptors or "nociceptors." Skin, muscle, and several other tissues have at least two different types of nociceptors; typically, the nerve fiber of one type is ensheathed by a thin fatty layer called myelin, and one or more others have fibers that lack the myelin sheath.

Nociceptors have been linked to human pain by a convincing experiment. Microelectrodes can be inserted through a human subject's skin into a peripheral nerve and manipulated to record the activity and initiate signals from a single sensory nerve fiber. When the nerve fibers of nociceptors are excited, pain is evoked. When the nerve fibers of other types of sense organs are excited, other sensations are produced, but pain never results.

Much human pain associated with disease derives from subcutaneous tissues such as bone or viscera such as the heart. Relatively little has yet been learned about the relationship between pain and activity of the sense organs of the viscera. Although skin and muscle nociceptors have been soundly documented, the existence of visceral nociceptors has not. Existing information on the cutaneous and muscle nociceptors provides the basis for development of pain-controlling drugs; comparable knowledge about the origins of visceral pain is clearly needed for the heart, intestine, blood vessels, etc.

Peripheral chemical intermediaries The events typically causing pain have as a common basis tissue injury or, in pathological circumstances, changes in tissue of the type produced by the inflammation associated with infection. Inflammation is a complex event involving increases in blood flow, increases in the permeability of blood vessels, and the interaction of certain blood cells with invading organisms or injured tissue. Many chemical substances are known or suspected to be intermediaries in the inflammatory process. These chemicals include amines, peptides, and derivatives of lipids.

Partly because pain is such a frequent concomitant of inflammation and partly from independent evidence, chemical substances have long been suspected to be intermediary between tissue events and activation of the sense organs that evoke pain. A connection probably exists between the kind of chemical events associated with the inflammatory process and either activation of the sense organs for pain or modification of their responsiveness.

For example, arachidonic acid, a breakdown product of a lipid cell constituent, and its metabolites, the prostaglandins and the leukotrienes, appear to play parts in inflammation and allergic reactions. Arachidonic acid and some of its derivatives excite the nociceptors. Furthermore, certain anti-inflammatory agents, such as aspirin, act at least in part by inhibiting the metabolism of arachidonic acid to prostaglandins. Aspirin-like drugs are effective against pain only when inflammation or vascular changes (e.g., headache) are present. Thus, it is possible that prostaglandins are mediators for nociceptor activation only in the presence of inflammation and that other, as yet unidentified, chemical intermediaries are involved in noninflammatory pain.

Testing by chemical agents has revealed important functional differences between the different types of pain receptors in a given tissue. For instance, one type is readily excited by chemical substances that evoke pain (e.g., bradykinin and capsaicin), while another type is unresponsive to the same agents. Such differences imply variations in either the transduction process or accessibility of the sensory terminals to various noxious stimuli. Another functional difference appears after certain kinds of tissue injury or in the presence of inflammation; the sensitivity of some nociceptors may be markedly enhanced while others are relatively unaffected. In fact, once excited, some nociceptors tend to produce persistent activity even though no further stimuli are evident; others do not.

Drug development The recent definitive establishment of nociceptors and the proof of their relationship to pain sensation have provided a far better focus for studies of drug action than has been available in the past. The fact that nociceptors of different types exist and that they have differing chemical sensitivities help explain some of the problems that have existed in drug therapy aimed at peripheral causes of pain.

Pain of peripheral origin, such as that associated with rheumatoid arthritis, is a major health problem. Heretofore, major pharmaceutical efforts to control pain in the periphery have been directed at developing nonsteroidal agents that interfere with prostaglandin formation from arachidonic acid. However, neither traditional drugs such as aspirin nor more recently developed nonsteroidal anti-inflammatory drugs that act on arachidonic acid metabolism are universally effective in suppressing pain of peripheral origin.

There is groundwork now for better-focused research on the basic molecular processes involved in activation of the nociceptors and their relationship to inflammatory damage. Investigation in this direction offers exciting new possibilities for developing drugs of either broader or special effectiveness for refractory clinical pain.

THE PERIPHERAL TRANSMISSION OF PAIN ACTIVITY

Peripheral nerves are composed of the many nerve cell fibers that convey impulses to and from the central nervous system. The nerve impulses are the result of a sequential opening and closing of channels spaced along the membrane that delimits each nerve fiber. When open, the channels allow ions to pass rapidly across the otherwise relatively impermeable membrane; this ion passage is part of the electrochemical event that makes up the impulse. The channels are complex entities consisting of protein and other organic components inserted into the lipid structure of the cell membrane. Blocking the channels blocks conduction in the nerve fiber and effectively disrupts communication. In the case of peripheral nerves, blocking the sensory (or afferent) fibers would cause anesthesia of the part of the body served by the nerve.

One proven diagnostic and therapeutic maneuver in treating certain types of peripheral pain consists of blocking particular nerve trunks with a local anesthetic, such as procaine. Local anesthetics generally work by blocking membrane channels. It has become apparent in recent years that the channels for different classes of sensory fibers may not be the same. Specifically, a number of the thin fibers of the type associated with nociceptors appear to have structurally different channels from those associated with the larger diameter sensory fibers. (The latter conduct the impulses producing bodily sensations of touch, pressure, vibration, and position.) Selective long-term blocking of the thin sensory fibers associated with pain would be an important therapeutic tool for the abnormal pain states often seen after peripheral nerve damage. It would enable other sensations from that part of the body while halting pain. The emerging knowledge about membrane channels and their molecular configuration offers promise for development of agents producing selective and persisting block of nerve conduction.

THE CHEMICAL MESSAGES AT NERVE JUNCTIONS

Study of the chemical transmission between nerve cells is one of the most active areas of neurobiological research. New work has shown that chemical interactions are part of the underlying organizational arrangements in the brain. Much research has been directed at determining how the activity in the periphery is transmitted to the cells of the spinal cord and brain and how it is modified by chemical agents impinging upon neurons of the pain pathways. Drugs working on the nervous system commonly act at neural synapses. This action usually consists of either synergy with a natural synaptic mediator or antagonism of its action. Thus the new knowledge of synaptic mediators and the characteristics of postsynaptic receptors has provided major opportunities not only for understanding the neural mechanisms basic to pain but also for targeting the development of specific drugs.

For example, there has been an explosion of information about neurally active peptide mediators (neuropeptides) and the existence of different, specialized receptor molecules on postsynaptic cells as the crucial link in their action. The neuropeptide story began as a general question about chemical mediators between sensory fibers and central neurons. It quickly shifted in emphasis when high concentrations of neuropeptides were found in the small cells associated with fine sensory fibers and in a particular gelatinous part of the spinal cord (*substantia gelatinosa*). The latter had long been suspected to be the termination point for the fine primary afferent fibers and to be related in some way to pain.

Some commentators and current textbooks erroneously imply that when a neuropeptide is present in a neuron it represents the synaptic mediator used by that neuron. An extension of this idea is that a primary afferent neuron that contained a particular neuropeptide employed the peptide for synaptic transfer to neurons of the central nervous system. On this basis and other indirect evidence, the peptide "substance P" was hailed by some a few years ago as *the* synaptic mediator for at least some pain afferent fibers. This interpretation began to change when it was found that other chemical mediators apparently were present in many neurons that contained neuropeptides. Furthermore, the experimental application of substance P and other neuropeptides produced slower-to-occur and longer-lasting effects than expected for the usual synaptic action.

Reconsideration of peptide action at synaptic junctions has produced a new and important concept. It is now proposed that some peptides with excitatory action function as long-term modulators of neuronal excitability and that the short-term signaling is done by a different chemical transmitter (cotransmitter) liberated by the same synaptic terminal. Several of the neuropeptides also act as hormones, that is, serve as messengers delivered by the bloodstream. Hormonal action fits with the idea of neuropeptides causing longer-term changes in neuronal activity.

Still another aspect of this story is the possible peripheral action of peptides. The sensory neurons with peptides in their central terminals also contain them in their peripheral fibers. Some of the neuropeptides discovered at central synaptic terminals have potent effects on peripheral tissue, including changes of the type seen with inflammation. The part that peptides released by sensory nerve cells play in the inflammation seen after tissue damage or infection needs to be established.

What has been learned about synaptic mediators, the peptides, and their possible cotransmitters is only a first step. We have yet to define the specific actions of many of the peptides and to determine how the receptor molecules they bind to then produce alterations in postsynaptic cells. If peptides are a source of the long-lasting effects produced in the central nervous system in association with pain, modulating their actions with drugs that block the peptide receptors is another avenue to new and crucially needed therapies.

Capsaicin, the substance that makes red peppers hot Part of the peptide history is linked to the ingredient of red peppers (capsaicin) that gives them their irritant qualities. Injected capsaicin both causes pain and, when repeated, has a remarkable desensitizing effect upon cutaneous inflammation and pain-like reactions in animals. A few years ago, it was discovered that capsaicin acts more or less selectively upon sensory neurons with unmyelinated peripheral fibers, one class of which are nociceptors. When capsaicin is injected into young animals it causes permanent loss of many unmyelinated fibers in peripheral nerves and a reduction in the number of neurons containing neuropeptides.

The mechanism by which capsaicin produces these toxic and desensitizing effects is not known. However, there is reason to believe that it will be possible to dissociate them through alterations in capsaicin's molecular structure. If this is done, the activity of at least one set of sense organs responsible for pain possibly could be eliminated on a more or less permanent basis. This is still another place where targeted drug development with important clinical use would be possible from further knowledge of the chemical interplay between neurons.

OPIATES AND OPIATE RECEPTORS

Opiates have been used by people for millenia. They have profound effects upon physiologic states mediated by the nervous system, including changes in consciousness, perception, the sense of well-being, respiration, and upon various visceral functions. The most widely appreciated medical credentials of opiates are for controlling pain. The potential for abuse and addiction also is common knowledge. Addiction and habituation (loss of effectiveness upon repeated use) are major limitations in the medical use of opiates for alleviating pain. None of the available opiates that are particularly effective for the control of pain is free of the curses of addiction and habituation. However, synthetic manipulation of opiate molecules has shown that some dissociation is possible between the various effects of opiates on different neural mechanisms.

Opiates, their synergists, and antagonists bind selectively and specifically to a special class of receptors on cell membranes. Several types of opiate receptors have been defined in the central nervous system. Differing affinities for structural variants of opiates suggest differences for the different opiate receptors although their definitive architecture must still be determined. Studies using the binding of chemically defined opiate analogues have demonstrated that opiate receptors are differentially distributed in the brain. In some regions, such as the dorsal horn of the spinal cord and certain parts of the brain stem, opiate-related activity has been implicated in the modulation of pain pathways.

Uncovering the class of chemical mediators for which the opiate receptors are specialized was a major discovery. These natural opiate-like compounds, or endorphins, are proteins and peptide fragments of proteins and thus are chemically very different from the opiate alkaloids, a surprising observation when considering the apparent specificity of the receptors. Endorphins have differing distributions in the brain, and the several kinds of opiate receptors have different affinities for the various endorphins.

The means by which activation of opiate receptors produces biological effects has been only partially worked out. Opiates and their kin in many situations have suppressive or inhibitory actions. Certain endorphins appear to be the naturally occurring inhibitory synaptic mediators in certain neural pathways.[1] Opiate receptors are present in the membranes of certain of the primary sensory neurons, and opiates can block the release of excitatory transmitters from these neurons. The process underlying opiate action includes blocking the ion flux that is required for release of transmitters. In addition to membrane ionic mechanisms, opiates act upon cellular energy systems.

It is common dictum that opiates are analgesics, although many clinicians are aware that the usual doses of opiates do not abolish pain. Recent studies linking work on the sensory component of pain to opiate action have provided a partial explanation. Opiates such as morphine increase the pain threshold for certain kinds of stimuli but not for others. An increased threshold appears for stimuli that preferentially activate a subset of nociceptors with unmyelinated fibers but not for stimuli exciting myelinated fiber nociceptors. These observations give insight into why opiate or other drug management of pain has been so difficult. Because several types of nociceptors are capable of initiating pain, each possibly using a different chemical synaptic mechanism, a particular agent (e.g., morphine) will act only on a subset of nociceptor-activated neural pathways.

---

[1] The sensory nerve fibers from the periphery only make excitatory synapses with central cells; that is, their chemical messages tend to make the postsynaptic cell more active. In the central nervous system there are inhibitory synapses as well, in which the activity of one cell suppresses that in another. Inhibitory synapses differ structurally from excitatory ones. They too are usually chemical in nature, with the chemical agents and the postsynaptic receptors different from those for the excitatory junctions.

The work on opiate receptors, in addition to providing some outstanding scientific advances, has also led to new therapies. An example is a method of opiate use for pain control in terminal conditions of metastatic cancer. The discovery of opiate receptors in the spinal cord in association with the central terminations of the primary sensory fibers suggested direct application of opiates to the spinal region. In some cases, opiates now are infused through catheters placed into the spinal canal to provide relief from pain produced by cancerous growth below the level of infusion. This is accomplished without the usual deleterious effects upon respiration and consciousness because the drug is not reaching the upper spinal cord and brain.

Molecular chemistry and the research opportunities New knowledge of the molecular chemistry associated with nervous system action and the neural apparatus for pain constitutes great advances. But, once again, what has been learned is only an introduction. Activation of some and possibly all pain receptors by the peripheral events that cause pain involves a chemical interplay. The full scope of the chemical agents and just how each acts upon the transducer terminals needs to be learned. New classes of compounds that mediate nervous system action have been found, and there is evidence that some of them are associated with the transfer of information between the peripheral sense organs and central neurons. Yet to come is the linking of specific transmitter agents, such as a given peptide, to the nociceptor system and then the determination of the nature of the receptor molecules to which they bind. The opiate receptors are a subset of such molecules, and what has been learned about them gives promise that better analogues and antagonists can be developed. Thus, appreciation of the molecular chemistry of the nervous system elements associated with pain has opened the door for a new and more effective pharmacology. Needed at this juncture are additional basic biological study of the chemical compounds and their actions and definition of their receptors. In this work there is hope not only for pain control but also for effective strategies in managing drug addiction.

THE BRAIN PATHWAYS IN PAIN

Although it is has been known for over a century that a particular pathway in the spinal cord was crucial for pain to be felt upon peripheral stimulation, the origin and terminations of the pathway are recent discoveries. It also turns out that this pathway is only one part of the central connections from the peripheral nociceptors. Additional important progress in defining the neural substrate for pain has come from studying the brain regions to which the nociceptors send their signals.

A rough outline now exists of the neural circuitry for the transmission of information from the nociceptors. Neurons making up part of this system include those in the dorsal horn of the spinal cord, a part of the brain called the thalamus (a major subcortical center for sensory messages) and a particular region of the cerebral cortex. Severing the tract leading from the spinal cord to the thalamic region blocks perception of pain (and temperature) and has been one surgical intervention used to treat persisting pain. However, such chordotomies, when effective, often lose efficacy with time. The reason for this change is not known; it may involve previously unsuspected alterations of connections within the injured nervous system. Now that the importance of the cerebral cortex in pain perception has been established, the basis and significance of interactions between such regions and other brain areas are a fruitful field both for further basic investigation and for study of therapeutic possibilities.

The discovery of specific pain receptors, and of specific pathways within the central nervous system to convey the information they provide, represents a crucial development for understanding pain. With this framework, even with limited details, it is now possible to direct further studies on the workings of the neural networks, their chemical features, and the pharmacology of their manipulation. Until this specific neural machinery was established and some of its features were determined, a systematic approach to pain management was not possible.

Pain modulation by internal circuits of the brain Sensory systems in all mammals are organized so that inputs from different sources interact. A given central neuron has hundreds to thousands of synaptic contacts upon it from other neurons. Reinforcement or inhibition of one input by another can occur. These interactions are the basis of feature detection in vision and tonal recognition in audition. They occur for pain at the spinal and higher levels. Much of the underlying neural connectivity for the interactions and the functional significance of the different elements remains to be worked out. Again, these interactions and the integrations that they represent depend upon chemical mediators, membrane receptors for them, and the activation of intracellular messengers.

A series of pathways descending from higher regions in the brain terminate at points where transfer of information takes place in sensory systems. These descending pathways originate at several levels in the brain and can profoundly influence the transmission of activity related to pain and other sensory modalities. In both experimental animals and human beings, localized artificial activation of certain regions of the brain can abolish pain or pain-related behavior. In part these modulatory actions are produced through the release of particular chemical agents (peptides and other compounds). One site of such action is at the first synapse between the primary sensory neurons and neurons of the spinal cord. Some of the inhibitory effects are actually presynaptic, i.e., they occur on the primary fiber termination itself. One amazing feature of this arrangement is that such a descending inhibitory connection to pain pathways can be activated by opiates acting centrally. These descending modulatory connections are part of the organization that permits a definitive reaction in the face of more than one demand upon behavior. They undoubtedly underlie conscious as well as subconscious control of reactions other than perception and would seem to be a factor operating in manipulations such as biofeedback effects upon muscle contraction.

The circumstances under which the descending pathways are activated in normal life and the implications of the various chemical messengers they employ are not clear at this time. Nonetheless, knowledge of their existence offers another access to control of pain signals through the administration of selective pharmacological agents. What is clearly needed to realize these possibilities is a further unraveling of the neural circuitry, the specific roles of various chemical agents and their receptors, and the way these combine to produce suppression of activity.

Interaction between sensory inputs has been used to control pain by deliberately activating sensory sources and pathways. Electrical stimulation of the skin, nerves, and pathways of the spinal cord have been used with varying success in controlling and modulating persisting pain. Acupuncture is one of the more widely employed procedures involving interactions between sensory inputs, but continuing experience with acupuncture for pain has dampened an early enthusiasm for it as a specific treatment method. Whatever its true value may be, acupuncture is frequently misused as a means of therapy. Proper evaluation of acupuncture and other techniques of sensory interaction in the treatment of pain have been hampered by poor protocols in basic research and by inadequate measurements in clinical investigations and trials. This important area needs rigorous research in the future.

PAIN AS A BEHAVIOR

Pain is a form of behavior. Its report will be influenced by all of the factors that modify behavior. We learn that human beings have pain only if they report it. Information from animals is even less direct. Thus, in contrast to studying neural reactions to noxious stimulation, investigating the perception of pain, particularly in a clinical setting, becomes a study of behavior. However simple this concept may seem, recognition of it was slow to come and represents a major source of enlightenment in clinical management.

A patient's expectations—that is, anticipation or bias—greatly influences the report or reaction. In the clinical situation, curing or healing of the event initially leading to pain does not necessarily end the complaint about it. Persistence of the complaint in these situations does not necessarily mean that the patient is malingering.

Some reports of pain after healing may originate from behavior or other central nervous system activities, but this kind of pain may also have a peripheral basis. An injured tissue can heal in terms of gross repair of damage and disappearance of evident inflammatory reactions; however, pain may persist, perhaps in a modified fashion, because of changes in the sense organs and the central systems they engage. For example, the injury and its repair may alter the physical arrangements of the nerve terminals in tissue and make nociceptors easier to activate than in preinjury circumstances. This could cause pain as if noxious events were taking place even though stimuli or circumstances were innocuous. Objective ways of evaluating such situations are not available and need to be devised. Low back injury and longstanding pain thereafter in the face of apparent healing is a troublesome and important case in point.

Recent attention to mental factors that modify patient behavior in instances of persistent or repeated pain has led to the recognition of special behavioral patterns. "Pain behavior" is a common clinical observation and may profoundly alter the lives of individuals. Unfortunately, evaluation of the mental contribution to the clinical complaint of pain is rarely systematic, and even when done it relies upon imperfect tools. It is known that the issue is important, but protocols for proper evaluation and management are not commonly available. For example, preventive measures such as proper counseling at the time of first contact with the health delivery system could decrease or eliminate many adverse behavioral patterns associated with persistent pain. Systematic implementation of such procedures have not been evaluated to a degree appropriate to the great economic and social burden created by the many persons with "chronic" pain.

THE DILEMMA OF CLINICAL KNOWLEDGE

Information from human experience has been a considerable source of increased understanding of pain and its mechanisms. But there are major shortcomings in the availability and use of such information, partly because of inadequacies in the careful evaluation of the human sensory status. The traditional procedures for clinical sensory testing are time-consuming and poorly quantified. Human sensory evaluation in most clinics is still done using intruments of the nineteenth century. Knowledge from the laboratory on effective means for such evaluations is seldom implemented at the clinical level.

Physicians have had greater concern for loss of human motor function and the obvious incapacities it produces than for losses of sensory capacity. However, many of the most distressing syndromes associated with pain involve sensory deficits, and full information is often missing because of the lack of systematic documentation of the nature of the defects. Protocols and instruments for efficient, effective evaluation of the human sensory status are seriously needed. Insularity of the clinicians and investigators has also contributed to this lack of adequate clinical knowledge; those primarily concerned with neurological aspects or with behavioral problems often give short shrift to each others' domains. Collaboration between those analyzing the neural and the behavioral features in future research is of first importance.

Clinical evaluation of pain management demands careful, controlled trials. Clinical studies of pain therapies rarely are undertaken with appropriate experimental designs and methods of evaluation. In the type of cases represented by back pain, prospective studies are needed to determine the circumstances under which an acute problem of pain in certain patients becomes chronic, persistent, and intractable.

SUMMARY

Research over the past decade or so has provided major new insights on the neural basis of pain and, as a consequence, on certain principles of functional organization in the brain. Now there is a solid basis for well-targeted approaches to further questions on the neural organization and how it may be manipulated to provide effective clinical management of a complaint that represents a major economic and social problem. In particular, the groundwork has been laid for the development of new classes of drugs to control or modify pain when it becomes a health problem. Also, it is now understood that particularly intense or persistent or recurring pain can of itself produce important behavioral and psychic changes. These in turn complicate its management. Progress in understanding and evolving therapeutic approaches to such behavioral and psychic associations with pain have lagged far behind advances in the biological sphere.

RECOMMENDATIONS

The major progress in understanding the biology of pain has come from basic investigations in neurobiology and molecular biology. The classical methods of neurobiology, electrophysiology, and neurohistology have provided a framework. The marriage of these techniques and knowledge with the methods of molecular biology to probe the presence and interactions of complex molecules has been responsible for some of the more dramatic new insights. The advances have come not so much from support dedicated to uncovering the neurology of pain as from general support for research on the organization and functioning of nervous systems. The panel believes that the best way to learn more about the neural basis of pain is to continue to support basic neurobiological research as it has been supported for the past decade. Special encouragement should be given to those interdisciplinary approaches that join the potent tools of modern molecular biology to modern neurohistology and neurophysiology.

There is an important caveat in this recommendation. The operation of the nervous system and such reactions as pain are based not only on the activity of the individual elements and their chemistry but also on a circuitry that consists of the interconnections between a multitude of cells. We are far from understanding all essentials of the circuitry or even the nature of all of the elements making up the circuits. Such information cannot be obtained by working *in vitro*, that is, with isolated tissue or cultured cells. Investigating brain mechanisms requires living animals or viable fragments of the brains of animals. Furthermore, there are major differences in the organization of the nervous system related to pain in different species, and some of these are of crucial importance in extrapolating information to human beings from experiments in animals. In other words, certain essential answers can only come from experiments on mammals, and for some matters only from work on carnivores and primates. Certain other types of experiments on general features of the activity of individual nerve cells and the effects of chemicals upon a general class of neurons may be possible *in vitro* on cultured nerve cells. Clearly, both kinds of research must be fostered with the choice of the species or tissue dependent upon the nature of the question.

The situation for the clinical approach to pain management is less satisfactory than that for research on basic mechanisms. Although advances in the management of pain have taken place, and a better understanding of the approaches necessary has been gained, there are major shortcomings in the use of clinical material for gaining new knowledge about disorders of pain and in applying present understanding to pain management. Overall, the quality of clinical investigation directed at problems of pain is not adequate.

We feel clinical data gathering, clinical management, and evaluation of therapeutic procedures could most effectively be improved by providing clear demonstrations of the appropriate approaches. These could be generated in a limited number of places in this country where clinicians and investigators with the necessary skills and insights happen to coexist. One way to encourage such centers, to focus their energy upon the development of better clinical tools for testing and evaluation, and to improve clinical trials of surgical, pharmacological, and behavioral interventions would be to provide funds dedicated to operating "demonstration centers" for pain research. Five such units at appropriate academic centers across the country would provide a reasonable base for developing appropriate approaches and as sites for training. Good models for diagnostic evaluation and clinical testing of therapies and preventive strategies will be important steps in translating what is known into better care at lower cost.

Report of the Research Briefing Panel on Biotechnology in Agriculture

Research Briefing Panel on Biotechnology in Agriculture

C. Eugene Allen (*Co-chairman*), University of Minnesota, St. Paul
Charles J. Arntzen (*Co-chairman*), The Dupont Company, Wilmington, Del.
Donald C. Beitz, Iowa State University, Ames
Neal L. First, University of Wisconsin, Madison
Robert T. Fraley, Monsanto Company, St. Louis, Mo.
Timothy C. Hall, Texas A & M University, College Station
George H. Lacy, Virginia Polytechnic Institute and State University, Blacksburg
Ronald L. Phillips, University of Minnesota, St. Paul
Peter H. Quail, University of Wisconsin, Madison
Clarence A. Ryan, Jr., Washington State University, Pullman
Ronald D. Schultz, University of Wisconsin, Madison
James H. Tumlinson III, Agricultural Research Service, U.S. Department of Agriculture, Gainesville, Fla.
James Truman, University of Washington, Seattle
Olen C. Yoder, Cornell University, Ithaca, N.Y.

Staff

James E. Tavares, Staff Officer, Board on Agriculture
Carla Carlson, Consultant
Phyllis B. Moses, NRC Fellow
Aida Neel, Administrative Secretary
Allan R. Hoffman, *Executive Director,* Committee on Science, Engineering, and Public Policy

Report of the Research Briefing Panel on Biotechnology in Agriculture

SUMMARY

The accumulated knowledge of science and its application to agriculture have resulted in high-quality agricultural products in the United States and the lowest food costs to consumers of any country in the world. With the advent of biotechnology and the ability to study and manipulate the genes of organisms at the molecular level, limitless opportunities now exist to recombine genes directly and to dramatically increase the efficiency of agricultural production.

The biochemistry, physiology, development, and behavior of animals, plants, microbes, and pest organisms and the genes that control these traits are now more easily manipulated for the benefit of agriculture. Biotechnology enables scientists to study the properties of resistance to disease and to move resistance traits across current reproductive barriers into agriculturally important species. New and highly sensitive diagnostics and improved vaccines can be produced.

The rate of advance in accumulating new information about the basic life processes of organisms, however, is not keeping pace with the rate of refinement of the tools of biotechnology. The ability to study and manipulate organisms using the tools of biotechnology requires broad fundamental information about specific organisms. Much of this fundamental information is currently lacking. Acquisition of new knowledge will depend on an integration of technologies from a number of biological disciplines as well as increased interaction between biology and other sciences. The role of U.S. agriculture as a major positive component in the U.S.-world trade balance hinges on the ability to continue to increase the efficiency of agricultural production and remain competitive in international markets.

Advances in biotechnology that will lead to application and commercialization in agriculture can be greatly enhanced by strengthening basic scientific programs for the study of agriculturally relevant organisms and infusing these programs with the powerful molecular biological techniques that are now available. Special attention must be given to (1) funding and structure of programs, (2) attraction of high-quality investigators to agricultural research problems, (3) flexibility in the regulation of science, and (4) technical resources necessary to conduct high-quality research.

INTEGRATION

Biotechnology is science in transition. It is supported by the relatively rapid development of new techniques that in turn are being used to increase the rate of advance in many areas of science, notably in agriculture. However, the gears of progress in biotechnology have not yet meshed; the rate of advance in fundamental scientific understanding and that of technical capability are not yet equal. The transition to application and commercial development of products derived through biotechnology will be complete when scientists are no longer limited by a lack of fundamental knowledge in the basic biological sciences.

Biotechnology makes use of combined knowledge in basic sciences such as biochemistry, cell biology, immunology, and physiology that is enhanced by an improved understanding of genetics at the molecular level. It is the ability to manipulate cells or whole organisms—through techniques based in molecular biology and genetic engineering—toward a product or application. The use of molecular biological techniques not only evolves from basic knowledge but will further contribute to that knowledge.

In the excitement that has enveloped these newer techniques, many decision makers and scientists have focused on the components of biotechnology rather than on the science as a whole. In separating the techniques from the basic science, emphasis has fallen on the techniques, the tools of molecular biology that allow scientists to isolate a single gene from thousands and then transfer it to a new host. But the power of such techniques provides little advantage for application and commercial development if the basic biochemistry, biology, and genetics are not complete. The tools to isolate a gene can be employed only after a particular gene has been identified and characterized biochemically or genetically. Without the fundamental information on the biochemistry, biology, and genetics of an organism, there is little foundation for biotechnology. There is nothing to manipulate.

The constraint to further progress in biotechnology and its resulting economic benefits is lack of integration. The sophistication of the molecular biology component of biotechnology has outstripped the component defined by a knowledge of specific organisms. To effectively integrate biotechnology as a whole, interdisciplinary research that combines organismal and molecular biology must be encouraged.

The techniques of biotechnology should not be isolated in the organization of research laboratories, in the structure or funding of competitive grants or special programs, or in the minds of decision makers or scientists. The encouragement of interdisciplinary groups that combine the expertise of researchers in basic and applied science with that of molecular biologists is an effective mechanism to bring about integration of biotechnology with the rest of biological research.

Integration promises two important benefits:

1. The quality of biological research will be enhanced through contributions made by classical and molecular techniques, and
2. The rate of discovery will increase, resulting in an expansion of knowledge in the biological sciences and the more rapid development of innovative commercial applications to agriculture and to other segments of the U.S. economy.

RESEARCH OPPORTUNITIES

Molecular biology is becoming successfully integrated with biochemical, physiological, and genetic studies in a few notable research areas. Integration and, subsequently, advances toward practical application in agriculture are developing in these few areas only because the basic knowledge is more complete and permits ready application of the tools of biotechnology. In many research areas directly related to the life processes of agriculturally important species, biotechnology is severely limited by a lack of detailed basic information in biochemistry, physiology, and genetics.

One measure of the rate of progress in the science of molecular genetics is the number of genes that have been sequenced in the past two years. An earlier briefing report[1] noted that fewer than 10 plant genes had been sequenced by 1982. In mid-1984 the DNA sequences for 86 plant genes and over 1,100 mammalian genes were known. Researchers can now test for correlations between biological properties described at the molecular level and traits desired by the plant or animal breeder for enhanced food production. Significant correlations of DNA sequences with selected traits will lead to an increased use of molecular biological methods in breeding.

The following discussion covers two areas: (1) improvement of production efficiency and quality of crops, food animals, and agricultural products; and (2) protection of plants and animals from biological stress caused by pathogens, parasites, and arthropod pests. Examples illustrate some areas of recent progress, and highlight the potential for improvement and application in many areas of research.

QUALITY AND EFFICIENCY OF FOOD PRODUCTION

In crop production and forestry, plants can be made more efficient primary producers in ways ranging from net improvements in capturing sunlight for use in the synthesis of carbohydrates, lipids, and proteins to improvements in the quality and balance of these food reserves for animal and, ultimately, human nutrition. Increased efficiency in animal agriculture will come from an improved ability to regulate reproduction, growth, and development toward better feed utilization and increased meat, milk, egg, or wool production. The opportunities to manipulate microbes as symbionts in nitrogen and mineral uptake by plants or as fermenters in the animal digestive tract are great. Engineered microbes are a source of enzymes used to make food products, for example, rennin for making cheese. Microbes can also be used to add value and quality to agricultural by-products.

Agricultural Organisms

Plants Rapid advances have occurred in areas supported by a more complete fundamental knowledge of plant processes such as photosynthesis. For example, the molecular properties of the plant chloroplast are now beginning to be defined. This self-replicating organelle contains many of its own genes; the cell nucleus contains the remaining genes coding for the chloroplast. Thus, photosynthesis and other chloroplastic functions, including carbohydrate storage and mobilization, are under both nuclear and chloroplastic control. Gene transfer techniques are providing new opportunities for gaining an understanding of the organism and the development of key biological processes including photosynthesis and control of plant growth and development.

For example, the DNA sequences that regulate expression of photosynthetic genes are currently being studied. Researchers have successfully transferred a gene encoding the small subunit of the photosynthetic enzyme ribulose-1,5-bisphosphate carboxylase-oxygenase between unrelated plant species. Genes for the enzyme, transferred from peas into tobacco and petunia, are stably inherited and, very importantly, retain their normal patterns of regulation, by light, and expression in green leaves. Success with the ribulose-1,5-bisphosphate carboxylase-oxygenase gene demonstrates that transfer of genes for important traits between plant species is feasible.

---

[1] Report of the Research Briefing Panel on Agricultural Research. Washington, D.C.: National Academy Press, 1983.

A recent, related development is the cloning of the gene that codes for phytochrome, the major photoreceptor controlling plant growth and development. Phytochrome functions as a switch controlling gene expression in response to light, as in the control of the ribulose-1,5-bisphosphate carboxylase-oxygenase gene. Access to the phytochrome gene will provide opportunities to study the range of plant developmental steps from cell enlargement to flowering by manipulation of this central control switch.

Additional opportunities to improve the quality and efficiency of crop plants are evolving from the study of genes coding for seed storage proteins. Goals are to obtain increased levels of essential amino acids such as methionine in soybean and lysine in corn and to eliminate antinutritional peptides such as trypsin inhibitor and lectin in soybean. The practicality of engineering seed storage proteins was recently substantiated with the transfer of a bean seed protein gene into tobacco. Modified tobacco plants contained more than 2 percent bean protein as part of the tobacco seed storage proteins.

Animals Genetic engineering techniques can be applied to enhance many traits in food animals. Daily administration of mammalian growth hormones, called somatotropins, that are produced by genetically engineered bacteria are being used to study growth in sheep, swine, cattle, and poultry. In dairy cattle the administration of bovine somatotropin increases milk production by 15 to 40 percent and increases the efficiency of feed conversion to milk. This product of biotechnology is being developed for commercial use.

Copies of growth hormone genes have been inserted into swine; the effects are currently being evaluated. Fecundity genes have been isolated in sheep and offer a method to improve reproduction and reduce production costs by increasing the number of offspring in domesticated animals. Two genes regulating the synthesis of enzymes for wool production are being cloned and studied in Australia. Sheep carrying these introduced genes are expected to produce about 25 percent more wool, with little additional energy or cost to production.

Microbes Microbes can be genetically engineered to produce animal growth regulators, diagnostics, and therapeutic substances such as interleukins, interferons, and vaccines. Improvements in food production and processing can be made using genetically engineered microbes. For example, the increased and improved digestion of cellulose and lignin by microbes in the rumen and intestines of domestic animals would permit use of fibrous waste materials, including brush or wood pulp, as feed. Thus, meat and milk might be produced more efficiently, and in areas that are limited in availability of grains. More effective production of essential amino acids and vitamins by the intestinal flora in food animals can lead to more nutritionally balanced and economical diets.

Limitations

The examples cited above demonstrate the power and potential of biotechnology. Practical improvements, however, will depend on a more complete understanding of the biology of major crops and food animals. In plants, few of the enzymes important in plant growth and development have been isolated and characterized. Cell culture and regeneration techniques that support gene transfer methods have not been fully developed for many major crops, particularly corn, wheat, soybeans, and rice. The ability to manipulate agriculturally important plant traits is limited by lack of understanding of the molecular mechanisms underlying these traits and by inadequate molecular biological techniques for gene transfer and plant regeneration.

For animals, further research needed to bring potentially productive results from biotechnology into commercial use includes (1) *in vitro* culture and development of embryos from the one-cell stage, appropriate for gene transfer, to the multicellular stage, appropriate for implantation in surrogate mother animals; (2) regulation and enhancement of gene expression; and (3) adaptation of safe viral vector systems to transfer desirable animal genes into multicellular embryos.

Use of genetic engineering techniques in microbes is currently feasible, but requires continued research in gene isolation, transfer, and testing for efficacy. Further basic research is needed on those agriculturally useful microbes that differ significantly from the well-studied bacterium *Escherichia coli*.

PROTECTION AGAINST PEST ORGANISMS

Diseases caused by bacteria, viruses, or fungi, and damage caused by insects and nematodes are major factors limiting agricultural production. Understanding the relationships between the plant or animal host and these pests can lead to control and treatment. Great advances have been made in the last few years by the application of molecular biological techniques to disease diagnosis and control. However, further research is needed on the basic interactions between host and pest before the newer technologies can be practically applied to protect crops and food animals.

Agricultural Organisms

Animals Hybridoma technology is a dramatic example of the application of biotechnology to disease detection and treatment. The technology permits the production of large quantities of pure, specific antibody from B cells in the immune system against virtually any antigenic substance. Such a monoclonal antibody is already being marketed as an oral prophylactic for calf scours, an enteric disease caused by the bacterium *E. coli*. This common and often fatal infection is responsible for livestock losses of $50 million annually. Monoclonal antibody treatments should be developed to prevent enteric diseases of viral, bacterial, and parasitic etiology in other domesticated species.

In addition, hybridoma technology applied to T cells in the immune system is being used to prepare lymphokines such as interleukins and interferons that increase the immune response of an animal against disease. Lymphokines might provide a method to enhance the immune response of the neonate, the stage of highest susceptibility to infectious diseases and, thus, of greatest economic loss. Monoclonal antibodies or genetically engineered interferons and interleukins should also be developed to control respiratory diseases such as shipping fever, which in cattle alone cause losses in excess of $250 million annually.

Diagnostic tests based on monoclonal antibodies are already on the market for infectious diseases of companion and food-producing animals. Similar specific and highly sensitive diagnostic methods based on monoclonal antibodies or RNA-DNA hybridization are being extended to disease detection and will most likely become routine ways of monitoring for infectious agents as well as for contaminants in meat and milk supplies.

Although studies of the mouse and human systems contribute to an understanding of animal immunology, direct application of hybridoma technology-based diagnosis and prevention techniques to food animals requires a detailed understanding of the basic biology, biochemistry, immunology, and genetics of these animals. To exploit the potential for biotechnology in disease prevention, detection, and treatment, basic biological research on agricultural species and their unique pathogens must be increased.

Plants The opportunities to apply genetic engineering techniques in combating diseases in crop plants are enormous. However, few products are on the verge of commercial application. A significant effort in research on specific interactions between host and pathogen is required before detailed disease control schemes can be designed and developed.

Studies on the mechanism of action of herbicides and of genes specifying resistance to herbicides will increase the options for protecting crops. Genes for resistance to commercial herbicides such as glyphosate, sulfonylureas, and atrazine should soon extend the number of crops that can be protected against susceptible, competing weeds.

Likewise, plant genes that confer resistance to pathogens and other pests can be characterized and ultimately transferred between plant species. Some resistance genes are quickly and easily overcome by pathogens; others have remained effective for more than 60 years. With an understanding of the molecular mechanisms of resistance, scientists will be able to select hardier resistance traits.

Pest Organisms

Biotechnology has been successfully applied to the study and control of viruses and pest organisms. For example, a simple, efficient, and extremely sensitive RNA-DNA hybridization test is now available to identify viroid-free plant material for propagation. Viroids, which consist solely of a small RNA molecule, are responsible for major losses in agricultural crops including potatoes. Several other viral detection systems have been developed using this technique.

Pathogenic bacteria, fungi, nematodes, and protozoan and metozoan parasites can be detected by this technique or a related technique using monoclonal antibodies. Furthermore, recombinant DNA procedures promise to go beyond detection to prevention or cure of crop diseases. For example, RNA sequences complementary to vital messenger RNAs inactivate these messenger RNAs through RNA-RNA hybridization. This strategy could be used to disrupt gene expression in pathogens.

The convergence and exploitation of knowledge from plant pathology and molecular biology is illustrated by the study of the bacterial plant pathogen *Agrobacterium tumefaciens*. Pathogenesis is mediated by a tumor-inducing (Ti) plasmid from *A. tumefaciens* that is transferred into the genome of host plant cells to direct the synthesis of special metabolites needed by the bacteria. Researchers have now adapted this Ti plasmid as a vector for transferring other genes into plants.

As additional molecular mechanisms of pathogenesis become known, prediction of measures that would interfere with many disease processes should be possible. Further research, however, at both the molecular and organismal level of host-pathogen interactions is required. At present, only a few laboratories in the United States are committed to research on the molecular biology of microbial plant pathogens. There is a pronounced deficiency of information on fungi, the major cause of disease in plants, and molecular nematology.

The lack of information on pathogenic nematodes illustrates the importance of integrating basic biological research developments with agricultural research applications. The life cycle, genetics, and developmental biology of the nonparasitic nematode *Caenorhabditis elegans* are being described in extensive molecular detail by several groups that use the organism as a model system to study cell differentiation. The biological information and genetic probes developed for *C. elegans* provide a foundation for studying pathogenic nematodes, which are responsible for $10 billion annual damage to all crops in the United States.

Opportunities to apply genetic engineering approaches to develop new, effective and environmentally safe insect control strategies are promising. Potential advances, however, not only await the development of gene transfer techniques for pest and beneficial insects but also depend upon a better understanding of general insect biology, including sterilization, pheromones, neurobiology, behavior, and insecticide resistance.

For example, the problem of insecticide resistance, which has plagued insect control programs, can be turned to an advantage. Single genes known to confer insecticide resistance, if introduced into mass-reared sterile insects or insect predators such as the wasp *Trichograma*, can permit concurrent use of biological control methods with chemical control programs.

A better understanding of insect neurobiology—its chemistry, metabolism, genetics, molecular biology, and impact on insect development and behavioral patterns—is needed. Effective biotechnological strategies based on more complete neurobiological information could then be designed.

For example, evidence indicates that insect selection of a host plant or animal can be altered by a single gene mutation. Further, direct genetic manipulation of insect behavior should alter courtship and mating activities. In the laboratory the female corn earworm can be induced to produce the pheromone, or chemical sex attractant, of a related species, the tobacco budworm. This manipulation is based upon the known biochemistry of each insect's sex attractant and results in fruitless mating attempts. Additionally, a new brain peptide that regulates the production of pheromones in these and other species has been discovered. Further investigations of specific enzymes and neural peptides can lead to strategies for blocking reproduction by disrupting sexual recognition.

The ability to transfer genes into a variety of insect species will be extremely valuable for pursuing pest control strategies. Gene transfer in the laboratory fruit fly *Drosophila melanogaster* via the P-element vector, a DNA segment that can insert itself into the fly's genome, has been extremely successful. However, this technique, developed for the *Drosophila* laboratory model, has not yet been extended to other insect species of agronomic importance. The Caribbean fruit fly, for example, an agricultural pest indigenous to the southeastern United States, can be used as a model to assess gene transfer techniques developed for *Drosophila*. Techniques developed for control of the Caribbean fruit fly would be directly applicable to future outbreaks of the Mediterranean or Oriental fruit fly. Such an approach could be extended to other insect pest species and genera.

Limitations

Two major factors currently limit the application of biotechnology to disease diagnosis and treatment and control of insect pests in agriculture. First, basic scientific knowledge is lacking on many important biological and molecular mechanisms of disease causation in both animals and plants, and on the molecular biology of plant and animal pathogens and insect pests. Second, biotechnological techniques for gene transfer and expression have not been developed for many pathogenic organisms and are incomplete or suboptimal for others.

These limitations can be addressed by an integrated approach to agricultural science that builds upon the fundamental interconnections between biology and biotechnology. New findings about the biochemical, physiological, and molecular genetic processes of organisms will provide fundamental knowledge that can be applied to improve agricultural productivity. This knowledge is needed to improve the molecular biological techniques essential in improving plant and animal systems in agriculture. In turn, these techniques can be used to advance fundamental research.

ENHANCING PROGRESS IN BIOTECHNOLOGY

The scientific opportunities in agricultural biotechnology are great. Research in several areas—particularly disease diagnostics for plants and animals and therapeutics and vaccines for animal health—has carried basic science into commercialization. In most areas, however, biotechnology must first contribute to expanding the basic understanding of agricultural organisms and their life processes. Combined with this knowledge, use of the tools of molecular biology and genetic engineering can lead to more efficient agricultural production through development of more effective methods and improved varieties of plants and animals. Research advances can further result in spin-off technologies with applications in research and medicine.

Special attention to (1) funding and structure of programs, (2) attraction of high-quality investigators to agricultural research problems, (3) flexibility in the regulation of science, and (4) technical resources necessary to conduct high-quality research is needed to overcome many of the barriers that have slowed progress in agricultural biotechnology.

FUNDING AND STRUCTURE

Funding for biological research applicable to agriculture comes from programs in a number of federal agencies in addition to the U.S. Department of Agriculture (USDA). Major resources include programs in biological sciences at the National Science Foundation (NSF), general medicine at the National Institutes of Health (NIH), and biological energy at the Department of Energy (DOE). Current funding levels and patterns, however, are only modestly promoting the integration of newer molecular biotechnologies with basic biological research in agriculture.

Research programs in the basic sciences should be structured to encourage diversity in science and to support interdisciplinary work for the integration of organismal biology with molecular biology. The funding of this research by several agencies, each with a different focus, should be encouraged and expanded.

Diversity

Industry has been increasing its support of basic science through special grant and contract relationships with universities. However, it is likely that private support for university research has reached its peak. Most companies are now building their own internal basic research programs rather than diverting further funds outside to support university research.

Efforts are being made in the intramural research programs of the USDA Agricultural Research Service to foster interdisciplinary research that includes use of the molecular tools of biotechology. A recent National Research Council report[2] discusses these research approaches in detail.

The level of support for extramural basic research that applies to the agricultural sciences varies throughout the federal granting agencies. The size of grants awarded also varies. If one assumes an indirect cost rate of about 40 percent, a grant from NIH pays twice as much per year as the average grant from NSF or DOE and four times as much as the average grant awarded by the USDA competitive grants program. In addition, NIH awards grants for up to five years; the other agencies typically award grants for only two or three years.

Agriculturally important research competes successfully for funding in grant programs outside USDA. However, much of the research needed to advance biotechnology in agriculture falls beyond the mandate of these other agencies. The possibility of continued growth of the USDA competitive grants program is encouraging. But, if these increases are accompanied by a shift in other agencies away from funding basic research with agricultural significance, the effect could damage current advances in biotechnology in agriculture.

---

[2] New Directions for Biosciences Research in Agriculture: High Reward Opportunities. Board on Agriculture. National Research Council. Washington, D.C.: National Academy Press, 1985.

Support for basic biology should continue throughout all federal agencies. This structure assures that high-quality research will be conducted from a variety of perspectives.

Special Initiative

Competitive grants are currently the most effective method available to assure—through peer review—that high-quality research is being conducted in high-opportunity areas. The current size of grants awarded, however, is not effective for encouraging interdisciplinary work that combines basic and applied agricultural biology with molecular techniques.

A special short-term initiative for 5 to 10 years is needed to provide the initial momentum for integrating molecular techniques into research programs in various biological disciplines. Special program grants should be awarded at an attractive funding level to interdisciplinary groups of three to five principal investigators whose efforts encompass the fundamental biology of agricultural organisms and molecular biology. This special initiative in basic biological and agricultural research could be structured through NSF and USDA.

Such an initiative, in addition to current programs throughout a number of agencies, could increase in a relatively short time the acquisition of fundamental knowledge and accelerate the application of biotechnology to basic agricultural problems.

Setting Priorities

The application of the tools of biotechnology to agriculture is at an early stage. During this period of research and discovery, it is extremely difficult to predict those areas that might readily yield new and valuable information when challenged with new techniques. Because high-return areas for biotechnology in agriculture cannot currently be predicted, efforts must be made to broadly support research of the highest caliber. High-quality research will lead to more rapid advance in the biological sciences.

Peer review is the best mechanism available for selection of high-quality research. Grants should continue to be awarded through this mechanism; thus, priorities are established by quality rather than by predetermined category.

To reinforce the continuity and focus of biological research in agriculture, a full-time position should be established for a science director within the USDA Office of Grants and Program Systems. Dynamic scientific leadership is needed to establish an integrated research focus for this growing grants program and to ensure that emerging technologies will be used in research on agriculturally important organisms.

HUMAN RESOURCES

Advances in agricultural biotechnology depend primarily on the quality and number of trained scientists drawn to a career in this area of science. Progress will depend upon attracting superior researchers to agricultural research problems. Scientific education should be directed toward establishing a solid foundation in the basic sciences related to agriculture, coupled with an understanding of the power of molecular techniques. This foundation will be critical in establishing strong, competitive programs in U.S. agricultural production and support industries that will successfully meet the increasing challenges of world agriculture.

Training Grants

Training grants and fellowships, awarded on a competitive basis by NIH and NSF, have been effective in enhancing and sustaining research excellence in selected areas of science. The USDA, however, has not placed major emphasis on training young scientists. A study[3] of plant biology personnel and training conducted by the American Council on Education reported that land grant institutions train 70 percent of the postdoctorates and 80 percent of the graduate students in the plant sciences. The 210 universities surveyed also reported a major need for new faculty positions in molecular biology.

Predoctoral and postdoctoral programs in agricultural sciences that offer an integrated approach to training in molecular biology and basic sciences including biochemistry, physiology, and genetics should be expanded.

To equip young scientists with the expertise necessary to advance biotechnology in agriculture, five-year training grants that emphasize the integration of basic science and molecular techniques should be administered by USDA or another agency such as NSF.

Career Development

Outstanding young investigators can be attracted to research that bridges organismal and molecular biology through the availability of career development awards. Such awards would support an individual in an academic research and teaching position for five years at an attractive funding level. Career development awards could be fully funded through USDA, or similarly to the Presidential Young Investigator awards, through matching support by industry or private foundations.

REGULATION

Along with the excitement of scientific development and application in biotechnology comes the responsibility of assuring safety and efficacy. In regulation—as in all areas of biotechnology discussed in this report—the barrier to progress in research and application is lack of sufficient understanding of the organisms and their interrelationships.

Basic Research

Regulation can be a barrier to the development of industrial application, because the rules and requirements of regulation have not advanced as quickly as the technology. The opportunities created by biotechnology are accompanied by the need for fundamental research to define guidelines and improve risk assessment. Genetically engineered innovations for agriculture must compete in the marketplace against many cost-effective, established technologies. The required submission of an environmental impact statement prior to each field test—at an average cost to the innovator of approximately $300,000—inhibits application and development.

Targeting new funds toward basic research that provides necessary answers to regulatory questions would be cost-effective over time in advancing technology and protecting the environment.

Reassessment

The Cabinet Council Working Group on Biotechnology organized by the Office of Science and Technology Policy is coordinating the policies of the agencies involved in regulating biotechnology. The interagency working group has stated that current laws are adequate to regulate biotechnology and no new legislation is needed at this time. The working group has also recommended that a scientific advisory panel coordinate the evaluation of scientific methods that support regulatory decisions involving biotechnology.

It is important that regulatory agencies be encouraged to review guidelines and regulations frequently—as often as every 12 months—to ensure that they are compatible with the increasingly rapid advance of the sciences. Regulations that are not at a scientific level comparable with current knowledge and industrial development and that delay testing for relatively long periods could slow research progress in industry and academia and raise development costs to prohibitive levels.

---

[3] Andersen, C.J. Plant Biology Personnel and Training at Doctorate-Granting Institutions. Higher Education Panel Reports Number 62. Washington, D.C.: American Council on Education, 1984.

RESOURCES

The impressive power and sophistication of molecular genetic techniques are often matched by the cost of equipment necessary to carry out these techniques. Other resources may be abundant but require assessment or management. These include an array of genetic collections and information on the genes they contain.

Equipment and Facilities

The cost of furnishing a laboratory to carry out high-caliber biotechnology research is approximately $300,000 to $500,000. Equipment such as protein sequencers and DNA synthesizers is expensive but essential. It is also essential that research teams throughout the country maintain well-equipped laboratories to conduct integrated research. Consolidation of efforts in a center of excellence is not necessarily a guarantee of superior research. The award of grants for equipment might be the most effective way to provide for quality research from a variety of groups.

Germ Plasm Maintenance

Biotechnological methods combine genetic resources with molecular analysis. There is an increasing need for access to the fundamental and irreplaceable genetic information that exists in collections housed at universities and state, federal, and international institutions.

As new techniques have emerged for mapping genes, including isozyme mapping and restriction fragment length polymorphisms, the problem of interpretation and coordination of data has become more complex. The present ad hoc efforts to map and coordinate the cataloging of genes in agriculturally important species leaves researchers with incomplete data bases and no mechanism for uniting all available genetic and biochemical data.

Existing resources should be evaluated and organized so that the valuable information they contain can be easily accessed by researchers and used to coordinate research on characterizing the genes of an organism. In addition, management of collections should be reviewed, and the quality of preservation and storage methods should be assessed. If the genetic information of certain plants and animals is lost or destroyed, the opportunity to improve a species by genetic manipulation will also be lost.

Report of the Research Briefing Panel on Weather Prediction Technologies

Research Briefing Panel on Weather Prediction Technologies

Richard A. Anthes (*Chairman*), National Center for Atmospheric Research, Boulder, Colo.
Ferdinand Baer, University of Maryland, College Park
Charles L. Hosler, Jr., Pennsylvania State University, University Park
Cecil E. Leith, Jr., Lawrence Livermore National Laboratory, Livermore, Calif.
Robert A. Maddox, National Oceanic and Atmospheric Administration, Boulder, Colo.
Richard J. Reed, University of Washington, Seattle
Robert J. Serafin, National Center for Atmospheric Research, Boulder, Colo.
Verner E. Suomi, University of Wisconsin, Madison

Staff

John S. Perry, *Staff Director*, Board on Atmospheric Sciences and Climate
Peter Abel, *Staff Officer*, Board on Atmospheric Sciences and Climate
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

The panel acknowledges significant contributions to this report by Drs. Rit Carbone and Joseph Smagorinsky.

Report of the Research Briefing Panel on Weather Prediction Technologies

SUMMARY AND RECOMMENDATIONS

The time is right for a coordinated national effort to improve our scientific understanding and prediction of weather on time scales ranging from less than an hour to more than a month. There is a high level of interest in the United States and other countries in the many exciting scientific questions, such as the generation and decay of convective phenomena including thunderstorms, squall lines and tornadoes, and the interactions of these *mesoscale* phenomena with both larger and smaller scales. The tools—new instruments for remote sensing from the ground and space, powerful computers, and sophisticated theoretical models—exist now and will continue to advance. In addition to answering some of the most challenging questions facing science, this effort will produce large and immediate benefits to society at a low incremental cost, creating especially great potential for high-leverage investments.

To take advantage of the scientific opportunities afforded by the advances in technology, we recommend:

- *High-resolution data sets and the accompanying research to use these data effectively in sophisticated numerical models of the atmosphere* for advances in the understanding and prediction of mesoscale atmospheric systems. Three types of data sets are required:

  Continuous Doppler radar data from a network of radars, eventually covering the United States to provide quantitative data of severe convective storms and wind measurements in most precipitation-free planetary boundary layers. These measurements will improve our understanding of energy transports in the atmosphere prior to storm developments and the interactions among various scales of atmospheric motion, information essential to improvement of models and forecasts.

  Continuous upper-air wind data from an array of wind profilers (Doppler radar wind measurement systems), also eventually covering the United States. This system will produce observations, nearly continuous in time and up to a height of about 15 km, of the wind, the most important single parameter for numerical weather prediction, and enable major advances in the understanding of atmospheric dynamics and thermodynamics.

  Special high-resolution data sets from field programs, such as those to be acquired in the National STORM Program.[1] These research data sets are necessary for advancing scientific understanding of mesoscale atmospheric phenomena. This better understanding will lead to improved treatment of physical processes (such as surface fluxes of heat and water vapor, and interaction between clouds and radiation) in both mesoscale models used for short-range weather prediction *and* global-scale numerical models used for medium-range weather prediction.

---

[1] The National STORM (Stormscale Operational and Research Meteorology) Program, a proposed major program to improve our ability both to understand and to predict mesoscale weather events, is outlined in *The National STORM Program—A Call to Action*, University Corporation for Atmospheric Research, Boulder, Colorado, 34 pp.

- *A focused research effort toward studies of several known sources of errors in global forecast models.* These include the tendency for models to develop systematic errors that lead to a model climate different from the real climate, inaccurate forecasting of slowly varying long waves, and inadequate mathematical approximations of such physical processes as radiation and condensation. The necessary research should consist of an integrated effort involving experts in radiation theory, cloud physics and dynamics, boundary layer theory, and applied mathematics.
- *Continued access to advanced supercomputers for both operational prediction and research.* Such computers are essential in extending our theoretical understanding of complex phenomena such as tornadic thunderstorms, to process the enormous volume of meteorological data obtained by remote sensing systems and to run global and mesoscale models of increasing resolution and complexity.

MODERN WEATHER PREDICTION

The atmosphere is a turbulent fluid whose changes encompass an infinite spectrum of temporal and spatial scales. These changes are governed by physical laws, a fact recognized more than 80 years ago when the prediction of atmospheric motion was first formulated as an initial value problem. Since then, weather prediction has evolved from an art into a science, with the quest for improved prediction on all scales of time and space centering on the development of improved numerical models based on an increasing understanding of the physical and dynamical processes governing atmospheric behavior. Continuation of this evolutionary process depends crucially on the availability of two types of data: global observations that are the basis for operational weather forecasting and observations obtained from special field efforts for research purposes. Additional advances in weather prediction require a balanced program of fundamental research that encompasses theory, development, and testing of numerical models, specially designed field programs, and improvement in operational observing, data processing, and communication systems.

Because of the nature of the prediction problem, it is customary to divide forecasts into different types according to time range. We identify four types: nowcasting (zero to six hours), short-range (six hours to two days), medium-range (two days to two weeks), and extended-range and climate (beyond two weeks). Each type has somewhat different requirements; for example, nowcasting (which includes warnings of severe weather) relies primarily on regional observations and extrapolation, while numerical models and global observations of the atmosphere, oceans, and land surfaces play a dominant role in predictions over the longer time ranges. Advances in satellite and computer technology and better understanding of atmospheric dynamics have made it possible to explore the feasibility of extended-range prediction using dynamical models, and some efforts are under way in this direction. This briefing emphasizes opportunities in nowcasting and short- and medium-range predictions, recognizing that progress on these time scales will also contribute to the scientific basis for extended-range and climate prediction.

IMPACT OF IMPROVED WEATHER PREDICTION

Advances in prediction of weather on time scales ranging from less than an hour to more than two weeks are possible, and would provide immediate benefits to the nation. On the shortest time scales, for example, use of a Doppler radar at Denver's Stapleton International Airport during an eight-week experiment in the summer of 1984 conclusively demonstrated the capability to save fuel, property, and life. Aviation authorities determined that the fuel saved due to timely short-range forecasts of low-level winds and attendant adjustments in runway assignments amortized the total cost of the experiment in only two weeks. Even more significant, at least one potentially fatal air crash was avoided as a direct consequence of a microburst wind-shear warning.

The economic and societal impact of severe thunderstorms that carry destructive winds, hail, lightning, tornadoes, and flash floods to all parts of the United States is well known. Some 908 tornadoes occurred during 1984 in 44 of the contiguous United States, causing property damage of at least several billion dollars. In one day, March 28, 1984, an outbreak of more than 30 devastating tornadoes struck North and South Carolina, killing 57 and injuring more than 1,200 people, and causing hundreds of millions of dollars in property damage.

Convective storms are but one manifestation of nature's extremes over the United States. On November 10, 1975, the Great Lakes freighter, the *Edmund Fitzgerald*, fell victim to hurricane-force winds spinning into a rapidly deepening cyclone. This low-pressure area had moved away from the mountains as a relatively benign storm, but then, unexpectedly and spectacularly, intensified as it approached the Great Lakes. Similar cyclonic storms, spawned over the Pacific, in the lee of the Rockies, or off the Atlantic coast, often grow to hurricane-like intensities. Virtually all economic activities, including *agriculture, power, transportation, shipping, state and local government functions, and general commerce*, are adversely affected by these events, which bring heavy rains, flooding, drifting snows, ice storms, high winds, and intense thunderstorms nearly simultaneously to different parts of the United States.

TABLE 1 Scientific Issues in Mesoscale Meteorology

Predictability
Scale Interactions
Forced Circulations
Instabilities
Waves
Boundary Layer and Surface Processes
Air-Sea Interactions
Regional Climate and Climate Change
Severe Weather
Cloud Dynamics and Microphysics
Cloud-Radiation Interactions
Fronts
Atmospheric Chemistry and Air Quality

On even larger scales, seasonal droughts, heat waves, unusually wet seasons, general basin floods, and early and late freezes produce widespread adverse effects almost every year through portions of the United States. Operational temperature guidance forecasts during the extremely cold period of the winter of 1983–1984 were frequently 20 to 30 degrees too warm over large portions of the southern Plains. These poor forecasts had enormously adverse effects on the strategies for operating natural gas pipelines.

The atmospheric science community is ready to undertake intensive studies of the physics of mesoscale phenomena—which are important to all scales of atmospheric motion—focused initially on the central United States, as part of the National STORM Program. The scientific issues to be investigated in the STORM Program are extremely broad, as indicated in Table 1. Research based on new technology will undoubtedly lead to improved understanding of the relationships between mesoscale atmospheric systems and their *larger-* and *smaller-*scale environments. Better understanding of the physics of mesoscale phenomena and the "scale interactions" is required if we are to improve weather forecasting on all four time ranges.

SCIENTIFIC BASIS FOR MODERN WEATHER PREDICTION

Development of mathematical models of the atmosphere began in the late 1940s and has advanced in parallel with the development of digital computers. Because of limitations in computer capacity and in the ability to model complex physical processes, the first primitive models dealt with only the larger space scales and neglected the effects of energy sources and sinks. Even so, these early models showed skill comparable to or slightly better than subjective forecasting methods for periods of one to two days. Since then, improved observations, coupled with enormous advances in computer capability, have allowed a variety of physical processes to be incorporated in modern models. As a result, the range of useful global forecasts has improved to at least six days.

Scientists are now developing sophisticated models of much smaller scale atmospheric phenomena. These models are providing new insights into the physical processes that govern such phenomena, laying the foundation for operational weather prediction on these smaller scales.

Medium-Range Prediction

The atmosphere is the prototypical chaotic nonlinear system. This was shown by the simplest atmospheric model, devised by Lorenz over 20 years ago, the starting point for modern mathematical studies of such systems. Because the atmosphere is chaotic, atmospheric models are sensitive to small variations in initial conditions and possess an inherent growth of error. These properties impose a theoretical limit on the range of deterministic predictions of large-scale flow patterns of about two weeks.

Prediction of the global atmosphere by the best computer models is currently accurate enough to be useful for a period of about a week, half the theoretical limit. A decade ago, the useful period was about a third of the theoretical limit. The improvement has come from a better understanding of atmospheric dynamics, models of higher resolution and accuracy, and better techniques for using meteorological data, especially satellite data. These advances, in turn, were made possible by faster computers with larger memories and by the accomplishments of the Global Atmospheric Research Program.

The improvement in medium-range weather predictions over the past decades is best exemplified by the experience at the European Centre for Medium Range Weather Forecasts (ECMWF), currently regarded as providing the most accurate medium-range weather forecasts. ECMWF's success is based on (1) having a very well-focused goal of improving medium-range forecasts, (2) access to the most advanced scientific computers, (3) emphasis on the scientific basis for numerical weather prediction, and (4) an effort to incorporate the best ideas and directly involve leading scientists from many countries, *including the United States.*

Short-Range Prediction

The interaction of all scales in the atmosphere is uncontested. While it is impossible, even with today's supercomputers, for a single model to resolve energetically active mesoscale circulations while simultaneously covering the global atmosphere, there is an increasingly important synergistic relationship between high-resolution mesoscale models that cover limited regions and lower-resolution models that cover the entire earth. The mesoscale models provide important guidance for time periods of about two days into the representation of small-scale physical effects (such as convective clouds) in the global models. Conversely, the global models provide essential meteorological information for the boundaries of the mesoscale, limited-area models.

Further advances in both global and mesoscale models have the potential to improve predictive skill significantly on all scales over the next decade. These advances require *enhancements in quality and resolution of data* together with *more accurate representation of physical processes* such as clouds, coupled with radiation, surface friction, and orography *in numerical models.*

Nowcasting

Predictions over the shortest time scale of zero to six hours include warnings of imminent severe weather and hence have great value in saving lives and averting property loss. Forecasts on this time scale are based on detailed observations and extrapolation of existing weather phenomena, rather than numerical models. Crucial to improved predictions and warnings are better observations from advanced radars and geostationary satellites. These systems provide temporally and spatially continuous observations of severe weather phenomena that almost always are missed by the conventional surface and upper-air observational network.

Although detailed observations from radars and satellites are necessary for significant advances in the prediction of atmospheric systems on short time scales, they are not sufficient. Advanced computers and broad-band communication systems are required to *collect, process,* and *assimilate* the data. In addition, further basic understanding of the structure and life cycle of mesoscale systems is required to enable scientists to *interpret* the wealth of complex information provided by the remote-sensing systems.

CURRENT LIMITATIONS TO SKILL IN WEATHER PREDICITON

Progress in medium-range prediction is limited by major gaps in the global data base, particularly over the oceans; short-range prediction of mesoscale weather over the United States is limited by the inadequate spatial and temporal resolution of the operational observing system. Prediction over all time ranges is limited by inaccurate and incomplete treatment of physical processes in numerical models.

Deficiencies in Data

In spite of modern meteorological satellites, the atmosphere over the world's oceans is still inadequately observed. Forecasts over the United States beyond about a day are adversely affected by errors in the observations made over the oceans. Moreover, data deficiencies are an important cause of the repeated failure of numerical models to predict rapidly developing intense ocean storms that pose a hazard to marine transportation and to coastal installations and activities.

Present oceanic data consist mainly of surface observations from ships and buoys, high-level observations from commercial aircraft, low- and high-level winds estimated from satellite tracking of cloud motions, and low-resolution vertical soundings derived from satellite measurements of outgoing radiation. These components of the ocean observing system fail to provide adequate horizontal coverage and provide almost no data at middle levels in the atmosphere.

Substantial improvement in ocean observations can be made both by expanding the capabilities of current systems and by further development of satellite technologies that have shown promise in preliminary testing. Routine ship and aircraft observations can be increased. Automated balloon-borne instruments, launched from commercial ships and dropsondes from aircraft, can provide needed thermodynamic and wind data. This low-cost concept has already been demonstrated in the North Pacific on Japanese car-carrying ships in cooperation with the Canadian Meteorological Service. Buoys that sense temperature, pressure, and winds can be deployed in data-sparse areas.

The SEASAT satellite, which operated briefly in late 1979, demonstrated the possibility for important advances in observation and analysis over oceanic regions. Particularly impressive were surface wind analyses based on scatterometer data. In one instance, the scatterometer winds (unfortunately not obtained in real time) revealed the existence of an unpredicted, intense storm that battered the liner *Queen Elizabeth II* and sank the dragger *Captain Cosmos*. In addition to wind observations, the SEASAT yielded passive microwave measurements from which water vapor, cloud water, and precipitation in oceanic storms were determined. These variables add a new and promising dimension to ocean analysis.

The required expansion of ship, buoy, aircraft, and satellite observations over the oceans can be achieved only through the cooperative efforts and mutual contributions of the meteorologically advanced nations. By pushing ahead vigorously with its own efforts, the United States can provide important international leadership in developing a substantially improved observational system.

Although the observational network over the continental United States is sufficient for medium-range global weather prediction, it is inadequate in both time and space for detection and prediction of mesoscale weather events. Upper-air data (pressure, temperature, humidity, and winds) are gathered only every 12 hours at spacings of 400 to 500 km (Figure 1). A staggering range of important weather phenomena, from individual tornadic thunderstorms to narrow bands of freezing rain, cannot be detected, much less resolved, by such widely spaced observations.

Until recently, it has been impossible to obtain observations, even over densely populated land areas, with the horizontal and temporal resolution necessary to detect and resolve mesoscale weather phenomena. And even if sufficient observations were available, limitations in communications and in computer power have prevented effective interpretation of the data and their use in mesoscale numerical models. Now, however, with the availability of remote sensing instrumentation, especially satellites and radars, together with the emergence of satellite communications systems and supercomputers, these technical barriers no longer exist, and the science of mesoscale weather prediction is poised for major advances.

DEFICIENCIES IN MODEL PHYSICS

Inadequate treatment of physical processes (such as clouds, radiation, and energy transformations at the earth's surface) in numerical models is a major source of forecast error. Several aspects of these model physics appear particularly important, including *climate drift, energy transfers at the earth's surface, and the dynamics of long waves and blocking patterns*.

Since global models start from observations of the real atmosphere, averages computed over many initial states reproduce well the real climate. But if averages are computed over many 10-day forecasts, one finds instead a model climate differing significantly from the real one. For example, many models develop average temperatures in high latitudes that are lower than observed, leading to a cold bias which results in jet streams that are too fast. This drift of models from the climate of the atmosphere is a source of error in all present-day model predictions.

Prediction models produce incorrect climates because they are missing or misrepresenting important physical forcing mechanisms. Leading components requiring improvement are the treatment of precipitation and associated latent heat release, the interaction of the flow with mountain ranges and the interactions of clouds and radiation. If the *climate drift* could be removed, then the range of useful large-scale forecasts would

Figure 1 This satellite photograph, taken on 7 June 1982, at 6 a.m. Central Standard Time, depicts a mesoscale complex of thunderstorms covering parts of Missouri, Iowa, and Illinois (the dark, nearly circular cloud pattern over this region). This mesoscale system caused more than 100 million dollars in wind damage. Another mesoscale convective system, a line of thunderstorms, is visible as an elongated dark band over northwest Texas and parts of Oklahoma and Kansas. These common mesoscale systems are inadequately observed by the operational radiosonde network, indicated on the figure by cross marks. The 440 km indicates the distance between the Boothville, Louisiana, and the Apalachicola, Florida, radiosonde sites. SOURCE: National Oceanic and Atmospheric Administration.

be extended by about one day, an improvement of about 20 percent.

There is increasing evidence of the importance of fluxes of heat and water vapor at the earth's surface in modulating atmospheric circulations on all time and space scales, and in particular, in medium-range weather prediction models. These fluxes depend in a complex way on the characteristics of the surface, including the surface reflectivity and roughness, amount and type of vegetation, and most importantly, the availability of surface moisture. Better model treatment of these *surface processes* would undoubtedly lead to better medium-range predictions, and might well be the key to skillful extended-range predictions of temperature and precipitation anomalies.

Another major source of medium-range forecast error is the inaccurate prediction of the slowly changing atmospheric *long waves* and the development, persistence, and breakdown of *blocking patterns*—cyclonic and anticyclonic circulations that persist for a week or more and play major roles in determining temperature and precipitation anomalies. The forecast errors in these large-scale waves are not easily explained, despite the fact that simple theories suggest that the longest waves should have the greatest predictability. Recent theoretical studies suggest that we are close to understanding the mechanisms that determine the evolution of these waves and may be able to improve their treatment in numerical models.

HIGH-LEVERAGE OPPORTUNITIES

Several scientific and technological advances are now converging, stimulating a growing emphasis on mesoscale weather research and forecasting. New remote sensing systems permit observations of the motion of the atmosphere and its heat and water vapor content at much higher spatial and temporal resolutions and with far greater accuracies than previously possible. Advances in computing allow us to consider mathematical modeling of mesoscale weather phenomena for research and operational prediction. Rapidly escalating capabilities in broadband communications technology (including, as a simple example, the transmission of weather forecasts and warnings via cable television), in addition to advances in data processing and display, facilitate the communication of critical data to users.

Implementation of New Technology

Computing has progressed in three important aspects. Large, fast mainframes permit the execution of numerical models that more nearly simulate the physics of the atmosphere. Microprocessor technology permits cost-effective, highly distributed interactive computing essential to the acquisition and real-time processing of complex and diverse data sets. Networking of various computers, through high-speed digital communications, allows distribution of data and analyses for control of large field experiments as well as physical interpretation of model output. Research on both global and mesoscale numerical models requires access to the latest, most powerful computers.

Electromagnetic remote sensing from satellites and from airborne, shipborne, and ground-based platforms permits four-dimensional observations on time and space scales appropriate for both short- and medium-range predictions. These observations span the electromagnetic spectrum from optical to radio wavelengths. The remote sensors most ready for application to prediction are Doppler microwave radars and Doppler UHF/VHF radars for the measurement of winds at all levels of the troposphere. As shown in Figure 2, the UHF/VHF radar wind profilers provide vertical soundings of the horizontal wind every hour, more than 10 times the frequency of operational radiosonde releases. Passive satellite radiometry at infrared and microwave wavelengths will continue to play a vital role for medium- and long-term predictions particularly over data-sparse oceanic regions, as demonstrated by SEASAT. Combined active and passive microwave sensors aboard satellites will greatly enhance our ability to define patterns of moisture, temperature, and surface winds over the ocean.

Some important observing system requirements cannot yet be met by means of remote sensing. Temperature and humidity soundings with high vertical resolution present particular challenges, but even here solutions are in sight. Highly automated, balloon-borne sensing systems utilizing Omega and Loran-C navigation aids can provide thermodynamic and wind profiles in real time over land and sea. The Global Positioning System, a worldwide navigation system based on satellites, promises even greater capabilities by permitting high-resolution profiles anywhere in the world. A still broader array of ground-based observing systems, complemented by research aircraft equipped with airborne Doppler radar, will be required as part of the next mesoscale field programs.

Field Programs to Provide Special, High-Resolution Data Sets

Special, high-resolution research data sets are urgently needed to document the life cycle of mesoscale weather systems and to initialize and verify new mesoscale research models. While many potential sources of data have been identified as part of the National STORM Program, a few are of particularly high priority because they offer the opportunity to function not only as state-of-the-art research tools but also as permanent observing systems in the operational prediction and warning service.

A *UHF/VHF Doppler wind profiler network,*

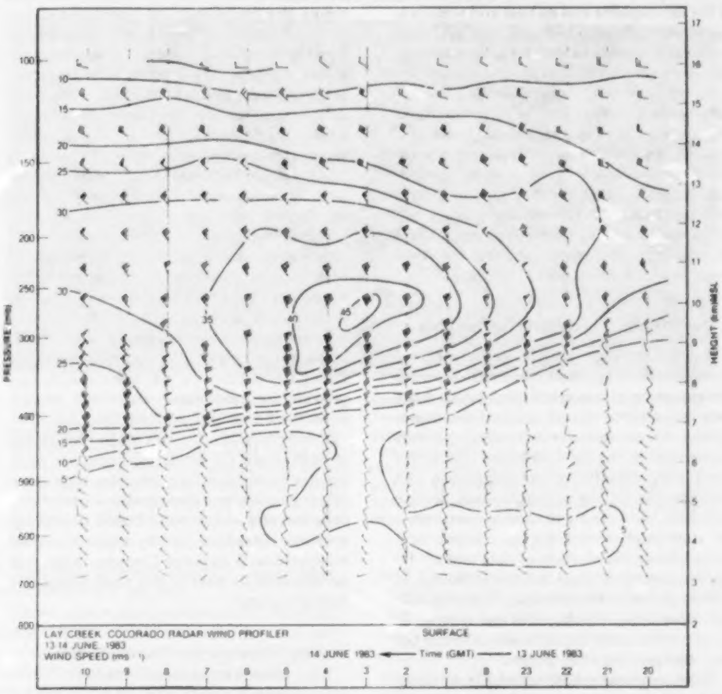


Figure 2 This figure depicts horizontal wind speeds above Lay Creek, Colorado during a 14-hour period in June 1983, as detected by a radar wind profiler. The line at midnight (0 hours) indicates the data provided by the one operational radiosonde launch that would normally occur in this time period. The observational detail provided by the profiler system far exceeds that of the radiosonde. Each symbol, representing one observation, gives wind speed and direction. Note that the jet maximum of $45 ms^{-1}$ (~90 mph) would not have been observed at all by the radiosonde. SOURCE: National Atmospheric and Oceanic Administration.

deployed over a large region of the central United States, would provide continuous observations from 1 to 15 km altitude of the wind, the single most important parameter needed for small-scale weather prediction (Figure 2).

A *microwave Doppler radar network*, initially deployed over the central United States, would measure low-level (0 to 1 km) winds in most circumstances, the complete wind field within precipitation regions of storms, and precipitation rates, thereby providing immediate operational benefits in the form of reliable, timely warnings of severe weather including flash floods and tornadoes. In addition, Doppler radars, located at airports, would detect low-level wind-shear hazards faced by military, commercial, and general aviation. These radar data would also contribute to the initialization and verification of mesoscale models for scientific research.

It is essential to build a high level of performance into the U.S. Doppler radar network. Highly sensitive 10-cm Doppler systems with polarization diversity and intelligent signal-processing capabilities would ensure quantitative estimates of precipitation in flash flood situations, hail detection, and the crucial measurement of low-level winds in the clear atmosphere. The NEXRAD design specifications, which have been carefully developed over the past decade, meet all of these requirements, with the exception of polarization diversity, which can be added to the system at a later date at a small incremental cost.

CONCLUDING REMARKS

The time is right for a coordinated national effort to improve weather prediction on all time scales, and particularly on the mesoscale. The interest in the United States and other countries in the exciting scientific questions is high. The tools—new observational technologies, powerful computers, and sophisticated models—exist now and will continue to advance. The potential economic benefit is high, the incremental cost low.

Report of the Research Briefing Panel on Ceramics and Ceramic Composites

Research Briefing Panel on Ceramics and Ceramic Composites

John B. Wachtman, Jr. (*Chairman*), Rutgers University, Piscataway, N.J.
William Rhodes (*Co-Chairman*), GTE Laboratories Inc., Waltham, Mass.
H. Kent Bowen, Massachusetts Institute of Technology, Cambridge
Donald Buckley, National Aeronautics and Space Administration-Lewis Research Center, Cleveland, Ohio
Eric Cross, Pennsylvania State University, University Park
Robert J. Eagan, Sandia National Laboratory, Albuquerque, N. Mex.
Anthony Evans, University of California, Berkeley
Gene Haertling, Motorola Inc., Albuquerque, N. Mex.
Joseph Panzarino, Norton Company, Worcester, Mass.
Dennis Readey, Ohio State University, Columbus
Roy Rice, W.R. Grace & Co., Columbia, Md.
Gerd M. Rosenblatt, Lawrence Berkeley Laboratory, Berkeley, Calif.

Staff

Don Groves, *Staff Scientist*, National Materials Advisory Board
Allan R. Hoffmann, *Executive Director*, Committee on Science, Engineering, and Public Policy

Louis R. McCreight, *Rapporteur*, The Aerospace Corporation, Los Angeles, Calif.
Robert Reynick, *Liaison Representative*, National Science Foundation, Washington, D.C.

Report of the Research Briefing Panel on Ceramics and Ceramic Composites

EXECUTIVE SUMMARY

Advanced or high-performance ceramics, now entering commercial use, are inorganic, nonmetallic materials having combinations of fine-scale microstructures, purity, complex compositions and crystal structures, and accurately controlled additives. Such materials require a level of processing science and engineering far beyond that used in making conventional ceramics. This is true for both areas of primary focus in this report: electronic ceramics and structural ceramics.

Our awareness of the tremendous potential that such materials can play in the U.S. economy and defense has been strongly reinforced by the work that has been carried out by other countries, notably Japan. Successful competition by U.S. industry depends upon economic factors and maintaining leadership in ceramics technology. Promising high-priority research opportunities can be grouped into four broad areas, each of which would support development in both electronic and structural ceramics:

1. New thin films and layer structures with improved properties.
2. Exploration of completely new, multicomponent ceramic crystal structures and composites.
3. The mechanical behavior of ceramics and tough composites.
4. Ceramic processing of large parts and assemblies.

To maintain leadership in ceramics technology the United States needs to strengthen engineering research on ceramics and to increase the number of properly trained scientific and engineering personnel for current and future work. Critical-mass efforts focused on ceramics but drawing upon advances in other disciplines, in advanced instrumentation, and upon computer-assisted modeling are needed.

Such efforts should be centered on ceramics, not dispersed in a variety of scientific and engineering disciplines. Still, other disciplines including chemistry, physics, chemical engineering, and electrical engineering need to be involved. Research should be done in teams of sufficient size and should have a common focus on major themes. This should augment, not displace, existing single-investigator programs in ceramics. The groups should have as a major goal the production of trained personnel. Industrial involvement is essential. However, these group efforts should not be an attempt to perform industrial development or to substitute for necessary, ongoing research and development activity in the industrial sector.

To achieve critical-mass efforts on ceramic research three mechanisms are recommended:

1. A modest number—a minimum of five—ceramic science and engineering programs should be established at universities over the next few years. Each such program should have a focus on aspects of ceramics in keeping with the base of faculty capability on which it is built, but the focus should not be too narrow, and there might be overlap between programs.
2. A summer institute program should be established to bring together leading ceramics researchers.
3. The feasibility of establishing at least one large interdisciplinary ceramics center be studied. Such a center could be equipped to carry out research on the basic science and engineering of ceramic processing and the application of ceramics on a practical scale. It could be viewed as a national facility for interdisciplinary experimental and theoretical work on high-performance ceramics.

INTRODUCTION

Ceramics and ceramic composites are a rapidly evolving category of materials capable of being tailored to have unique combinations of electrical, optical, mechanical, and chemical properties that make them essential and irreplaceable in many engineering applications. Conventional ceramics have both enabled and enhanced many important aspects of our modern technological society; the next generation of ceramics will play a further enabling role for still higher performance devices.

The utility of high-performance ceramics in a variety of electronic, optical, and structural applications is based on their strong chemical bonds and the mixed ionic-covalent character of the bonds. This bond character provides an electrical insulating and semiconducting behavior; optical transparency; ferroelectric, ferromagnetic, piezoelectric and pyroelectric behavior; high hardness and strength; and good corrosion and oxidation resistance.

The utility of advanced ceramics is based also on their ability to play both functional and structural roles, sometimes in combination and often under conditions of high temperatures and severe chemical environments. Functional roles include passive and active components in semiconductor devices, optical devices, motors, transducers, and sensors. Structural roles include applications in which high specific strength, high specific elastic moduli, high wear resistance, and high corrosion resistance or a combination of these properties are needed. Advanced ceramics are also finding increasing use in information processing (electronic and optical), in automation (input-output systems and computer components), in materials processing (including dies and cutting tools), and in engines.

Ceramics are an enabling technology. The competitive performance of many devices and large systems depends on ceramic components that currently make up a small but vital part of the total package. The competitive economic leverage of superior ceramic components is large. These superior components depend, however, on carrying the scientifically demonstrated performance of new ceramics into engineering realization as well as on pursuing the science of still better ceramics. The challenge of realization is a continuum of problems in advancing the science and engineering of ceramic processing to levels of purity, perfection, and scale not previously achieved.

Advanced ceramics is a critical arena of intense international competition. The United States at present is strong in the basic science of ceramics but appears to be increasingly threatened in the extension of this knowledge into the engineering and commercialization of advanced ceramics. Japan, in particular, has a 10-year program in high-performance ceramics with extensive industrial participation and commitment. Japanese firms already control 70 percent of the free-world market for electronic substrates and appear likely to dominate much of the rest of the market for advanced ceramics in the near future unless the United States responds immediately in a more effective manner. Recent events in Europe indicate that similar strong competition founded on excellent technical capabilities is also developing there.

The size of such overseas efforts, their coordinated approach by industry, university, and government, and the close coupling of research and development to applications all underscore the seriousness of the competitive race faced by the United States. Many U.S. firms not previously active in ceramics have identified this as a promising area and have started new programs in ceramics; many firms with existing ceramics programs have strengthened these and expanded their scope. Over 50 firms have joined university-industry programs in ceramics in the last three years. There is a strong commitment to succeed, but also an urgent need to expand the ceramics manpower base if these new initiatives are to prosper and bear fruit. Thus, the time is opportune for government, university, and industrial cooperation in high-technology ceramics.

The key issue discussed in this brief report is the urgent requirement for an increase in U.S. basic science and engineering research for high-performance ceramics. Further technological improvements will result rapidly from the proper application of resources aimed at strengthening and broadening various aspects of basic and applied research in the field. The application of such resources includes not only funding of candidate projects and programs but also directions to ensure a supply of trained professional scientists and engineers necessary to advance the technology.

Designers and manufacturers of advanced engineering systems in the electronics, transportation, computer, manufacturing, and communication fields are becoming increasingly aware that ceramic components offer improved system performance or function. The report focuses on two general categories in which opportunities and growth are expected: *electronic ceramics* and high-performance *structural ceramics*.

ELECTRONIC CERAMICS

At present, electronic applications dominate the world market for high-technology ceramics. Electronic ceramics are essential to the electronics industry, the base of America's current industrial strength. Major applications include low-loss optical fibers, which are causing a revolution in broad-band information transfer; multilayer ceramic-to-metal interconnecting and mounting packages for critical silicon semiconductor integrated circuits; ceramic multilayer chip capacitors (MLCs) of exceptional volumetric efficiency, required to decouple integrated circuits; piezoelectric ceramic transducers for sonar and medical ultrasonic tomography equipment; and chemical, mechanical, and thermal sensors for automobiles and automated manufacture and control.

Ceramics for electronic applications are high-band-gap, largely insulating ceramics and glasses in which a broad range of dielectric, elastic, optical, thermal, electrical conductivity, piezoelectric, and pyroelectric tensor properties can be manipulated and controlled to close tolerances. In most applications the ceramic does not stand alone but is bonded to a metal (for electrodes and conductors) or polymer, which requires precise control and understanding of the interface. Often the metal and ceramic parts are in closely spaced lamellar structures, as for multilayer interconnection packages or multilayer capacitors, and must be coprocessed through the whole ceramic forming and firing process to yield defect-free monolithic hermetic packages.

For integrated circuit (IC) packaging, glass-bonded alumina ceramics are widely used. The green ceramic is usually tape-cast with a suitable organic vehicle, and molybdenum metallization inks are screen-printed for the interconnection wiring. Automation is extensively used in inspection, punching, printing, and lamination. Precise control of shrinkage on firing is essential for the more complex packages. A high degree of perfection has been achieved in this technology. Future-generation IC packages will require ceramics with higher thermal conductivity as the size of the package shrinks to increase computation speed.

Ceramics with lower permittivity are urgently required for advanced electronics. One of the simplest ways of lowering the permittivity in a ceramic or glass is to introduce a void space with a relative permittivity of unity. There is a real need for new processes and design methodologies to introduce such void networks without serious deleterious effects on mechanical strength.

Current technology, with the exception of optical fibers, is generally an extension of standard ceramic technology—i.e., powders pressed or formed with binders and sintered to densify ceramics. Results of recent research on solution processing have improved the purity and size distribution of powders and have refined the microstructural control required to attain desirable electrical and mechanical properties. However, the present technology is crude relative to that used to process semiconductors.

Major processing advances should also center on developing the underlying science and technology to make lower-temperature pinhole-free films in the 1 to 25 micrometers thickness range. These are essential for the evolution of the next generation of multilayer packaging and multilayer capacitors. An evolutionary change to ceramic processing for packages using organically derived ceramic films has started. An intense interdisciplinary effort involving organic chemists, ceramists, metallurgists, and physicists is needed to understand and exploit this emerging technology. At present, existing ceramic systems (e.g., Al_2O_3) are being explored, but to meet requirements for low dielectric constant, high thermal conductivity, high strength, and matched thermal expansion to the semiconductor, new materials must be developed.

A revolutionary change in electronic ceramic processing will occur when semiconductor processing is applied to the ceramic components. Approaches to control the physical properties of ceramics on or near a chip, such as selective area ion implantation and laser-induced recrystallization, will greatly enhance packaging capabilities. Related studies to improve the crystalline perfection of thin-film structures will make possible the fabrication of high-permittivity ferroelectric and pinhole-free thin films for capacitor applications.

Barium titanate ($BaTiO_3$) is the dielectric in most current multilayer ceramic capacitors. Lead and bismuth borate glasses are frequently used to lower firing temperatures and to permit the use of less expensive silver and palladium alloys for internal electrodes. Tape-casting is widely used for green-forming. Precise processing is required because commercial dielectric bodies must often be fired to a specific nonequilibrium phase assemblage to achieve the permittivity and temperature characteristics required for application.

Electrically active ceramics (e.g., ferroelectrics and varistors) require precise control of microstructures and chemistry to attain suitable properties. Although the physics of varistor and ferroelectric behavior is not fully understood, performance has been significantly improved through process control. For example, the switching field for varistors can be increased to 120 kV/cm from the 5 kV/cm found in commercial devices by using ceramic precursors derived from solutions to control grain size and maintain a very fine-grained structure.

In *sensor applications*, electronic ceramics fulfill many different roles that make use of the unusual property combinations achievable. For force, stress, and vibration sensors, the polarized piezoelectric ceramics in the lead zirconate titanate (PZT) and related families and the newer electrostrictive materials in the lead magnesium niobate relaxer compositions are currently of major interest. In a range of chemical sensing applications, the semiconductor ceramics in the zirconia and titania material families and zinc oxide currently are very important. For oxygen-sensing and pressure control, zirconia (ZrO_2) and titania (TiO_2) are useful. Furthermore, in the titania- and zinc oxide-based systems, specificity to a number of hydrocarbon gases can be controlled. New advances and their integration with semiconductor technology ensures a bright, expanding future for these systems.

Sensors based on ceramic systems are in a dynamic research and engineering phase. The balance between sensitivity and selectivity needs to be addressed through fundamental studies in such areas as catalysis, as well as through clever engineering designs to build in film sensors of varying sensitivity or selectivity on one chip. The number of sensors in the automobile, medical, and home market will grow markedly, and many will be ceramic-based because of the inherent stability of this class of materials.

Current *optical fibers* are silicate glasses that contain only parts per billion levels of impurity. Precise control of the refractive index profile is achieved by the addition of germania in the preform. Processing, in this instance, has kept pace with theoretical development. Absorption coefficients are near their theoretical limits for silicate fibers, but research on longer-wavelength fluoride-base glass fiber is in its very early stages. Fluoride glass fibers are predicted to give an order of magnitude improvement in performance, but the knowledge base of fluoride glasses is still limited.

It has been suggested that the new ability to use optical fibers for high-density, high-bandwidth communication must eventually force a change from electronic to photonic signal processing. In this revolution, electronic ceramics may play a vital role in providing the materials base for the simpler hybrid technology that will be required before full integration is accomplished in gallium arsenide and related semiconductors. Research is needed to develop existing families of perovskite, tungsten-bronze, and lithium niobate electrooptic materials to the level where deposited or in diffused optical guiding structures can be reproducibly fabricated.

In a more sophisticated area, holographic signal processing can be achieved, using the photorefractive effect, in electronic ceramics such as $LiNbO_3$, $KNbO_3$, and $BaTiO_3$. This effect has application in interconnection control as well as more unusual possibilities in nonlinear coupling for two- and four-wave mixing and signal up-and-down frequency conversion.

The future promises continued dynamic growth in these areas, together with burgeoning new families of applications in hybrid optoelectronics that use ceramic components in optical waveguides; in electrooptic and acoustooptic modulators and switches; and in sophisticated nonlinear optical configurations for four-wave mixing, phase conjugation, and optical frequency multiplication. However, to realize the potential in these new areas it will be necessary to broaden the base of ferroelectric single crystals that can be made available in highly perfect single domain state. To move to more complex compositions so that properties can be adjusted to desired values one must learn new techniques for the growth of new perfect crystals in new solid solution compositions.

STRUCTURAL CERAMICS

High-performance structural ceramics uniquely combine high strength, strength retention at high temperatures, high hardness, dimensional stability, good corrosion and erosion behavior, low coefficient of friction, high elastic modulus, and low mass density. Structural ceramics are used as coating, monolithic, and composite components. Major applications include tooling for metal working; wear components in a variety of abrasive environments; bioceramics for bone or tooth replacement; and military ceramics for radomes and armor.

Major efforts are under way to apply structural ceramics to automotive reciprocating engines for wear components, turbochargers, and a variety of diesel engine components; gas turbines for regenerators, recuperators, and stationary and rotary components; and ceramic bearings for both rotary and stationary elements. Success in any of these future applications will dynamically affect market growth, improve energy efficiency and the balance of trade, and reduce dependency on critical metals. The combination of cleaner burning by gas turbines and greater fuel efficiency for any heat engine using ceramics would improve environmental air quality.

Monolithic structural ceramics are currently based on silicon nitride, silicon carbide, partially stabilized zirconium oxide, or alumina systems. Processing and microstructural advances have made significant strides in improving mechanical properties and reliability. Considerable research effort is still needed to refine further these parameters before there is wide-scale utilization of such systems in structural ceramics. The ambient temperature brittle fracture behavior is well understood from a fracture mechanics perspective. The strength-limiting flaws have been identified, and work is under way to either shrink their size or eliminate them through improved processing; this work must continue in order to enhance reliability. Improvements are needed in powder synthesis, powder properties, near-net-shape fabrication methods, microstructure control, mechanical properties, and nondestructive testing methods. High-temperature fast and slow failure mechanisms due both to mechanical and to chemical environments must be understood and controlled. It is possible that new alloys of existing matrix materials with improved properties can be developed. Major efforts are required to realize the full potential in long-term high-temperature applications.

The existing structural ceramics for high temperatures are very useful but offer a limited range of choices. At intermediate temperatures the range of choices is significantly broadened by the availability not only of other binary compounds but also of a variety of ternary and higher compounds. These latter are almost exclusively compounds of silica with various lighter metal oxides such as MgO or Al_2O_3. For thermal structural applications, lower stiffness and especially thermal expansion are key parameters to control for wider applicability. The extensive use of these silicates stems from the existence of most of them in nature but is also based on one or more properties or a combination of properties giving superior performance over binary oxides. This superiority stems from the greater range and complexity of crystal structures that commonly accompany a greater number and diversity of atomic constituents. The diversity of structures and resultant properties is not limited to thermal structural applications but may also find applications based on unique properties. New refractory oxynitride and oxycarbide glasses are under study, and these may broaden the range of properties available for monolithic glasses or as additions to existing materials for liquid-phase sintering.

The basic research components for identifying new, more complex ceramics are predictions of phase relations, crystal structures, and key properties of interest. The greatest need for estimation and guidance is in areas such as thermal expansion and high-temperature solid-to-solid phase transitions. Important criteria are identification of materials having (a) low expansion (e.g., carbosilicides and nitroborides) for improved thermal shock-stress resistance; (b) high expansion for coatings more compatible with metal substrate expansion; (c) high thermal expansion anisotropy for use in composites; (d) phase transformations useful for phase transformation toughening; (e) easily sheared structures for lubricity; (f) natural fibrous structures for toughness (e.g., more refractory analogs of jade); and (g) more refractory glasses for sintering aids, bonding, fibers, and bulk glass applications.

Structural ceramics for *wear and bearing applications* depend on an understanding of their fundamental tribological properties. Some measurements of adhesion, friction, and wear behavior primarily of monolithic ceramics have been conducted, and limited studies have been made of the effect of environmental constituents on tribological behavior. Surface studies have examined some interfacial phenomena for ceramics in solid state contact with other ceramics, with metals, and to a limited extent with polymers.

The fundamental tribological properties of ceramics must be related to other physical, mechanical, and chemical properties to provide a basis for material performance. Interface studies of ceramics in contact with themselves and other materials need to be expanded so that theoretical models can be developed to predict interfacial behavior.

There is little understanding of conventional fluids for the lubrication of ceramics. Liquid lubrication technology has been developed primarily and nearly exclusively for metallic systems. The mechanism of lubrication of ceramics must be examined.

Coatings, ion plating, sputter deposition, and chemical vapor deposition should be investigated for improving the tribological performance of ceramics. The interface between coating and substrate is critical because, under the high contact stresses encountered in mechanical components, failure at this site limits coating life. Surface treatment of monolithic parts or coatings by ion implantation, laser glazing, and electron-beam treatment should be investigated to improve the tribological performance of ceramics.

Thermal barrier coatings for applications such as adiabatic diesel and gas turbines are important and warrant further research and engineering to optimize long-term performance. New processes may be required to fabricate these coatings economically with the required microstructural perfection for long life. Little is known, for example, about the effect of microstructure and interfacial properties on high-temperature environmental and mechanical fatigue.

Ceramic matrices combined with particulates, whiskers, or fibers of a different ceramic compound or metal for enhanced performance have yielded *composites* with five times the resistance to fracture (toughness) of monolithic ceramics. While ceramic matrix composite research is embryonic, it has a bright, but uncharted, future. Composites could find application in lightweight, stiff, dimensionally stable space structural members, as key building blocks for new high-temperature rockets, or in many of the applications discussed earlier.

Ceramic matrix composites toughen by increasing the fracture energy of the material system over what can be achieved in monolithic ceramics. Toughening is achieved through crack deflection around the fiber (whisker or particulate), the increased stress necessary to break the fiber, or the energy required to pull the fiber from the surrounding matrix. Typically, a well-designed composite will fail noncatastrophically under the highest loads. The interfacial strength of the matrix-reinforcement phase is a key parameter to adjust for optimal toughening. Significant advancements in recrystallized glass reinforced with silicon carbide or carbon fibers have been made by controlling the interfacial bond strength. A number of other parameters are crucial in this field, such as thermodynamic stability, thermal expansion mismatch, relative elastic moduli, processing protocols, and microstructure.

New research is needed on quantifying the improved mechanical properties for composites, particularly fracture energy or toughness, as the fracture mechanics routines for monolithic ceramics are invalid. Literally thousands of combinations of matrix and reinforcement phase are possible. An empirical approach has produced some promising candidates, but this is not recommended as a sufficient approach for the future. Instead, a judicious selection of systems should be based on sound scientific principles, including thermodynamics, and relevant physical properties. It may be necessary to collect fundamental thermodynamic and phase diagram knowledge before selecting a high-temperature-compatible reinforcement phase. Once chosen, new fibers and/or whiskers may be required. Interfacial properties between phases need to be investigated and perhaps modified through coatings to adjust for the desired level of bonding. Processing composites to form dense, flaw-free, uniaxially aligned and, alternately, orthogonally aligned composite structures requires major science and engineering efforts. Quantitative microstructure characterization and iterative feedback through the synthesis, processing, and properties loop will lead to optimized structures that might revolutionize such systems as the gas turbine, rocket nozzle, or radome technologies.

DESIGNING FOR CERAMICS

Ceramics present special requirements for successful application under high stress. They characteristically undergo brittle failure and may show creep or slow crack propagation at high temperatures or under reactive conditions at stresses somewhat below the fracture stress. Stress concentrations can cause catastrophic failure. A new field of designing from brittle materials, including ceramics, is taking shape. Aspects include computer-assisted analysis of the stress distribution in great detail, probabilistic design based on statistics such as the Weibull statistics of failure, and time-dependence of mechanical properties of materials under load. Design should also take into account the capability of processing large and complex shapes and the impact on properties of scaling up to large sizes. Joining of ceramic parts into assemblies also presents special problems for both design and processing.

PROCESSING OF CERAMICS

The theme of ceramic processing runs through the discussion of both electronic and structural ceramics. Improvement in performance requires improvement in the science and engineering of ceramic processing. Promising directions include the use of ultrafine powders and the use of chemical routes that supplement or bypass some of the powder processing stages. Special processing requirements include the processing of fine-scale layer structures, processing of ceramic composites, joining of ceramic parts, and processing of large and complex parts to near net shape to reduce final machining requirements. Fundamental processing studies would greatly benefit from a connection with the engineering processing of large parts. Several mechanisms can occur in parallel in ceramic processing and may obey different scaling laws. In addition, large parts frequently have thin sections so that modeling and optimization of processing based on small-part fundamental studies need the test of comparison with experiments on large, complex, and costly parts.

Ceramic processing technology is relatively attractive from the point of view of raw materials supply and hazard control. Most of the required raw materials are domestically abundant. Solid ceramics in finished form are among the least hazardous materials.

Hazards associated with some powder, fibers, and vapors occurring during processing are regularly addressed by industrial safety engineers. Studies show such hazards can be controlled by proper process design.

EDUCATIONAL CONSIDERATIONS

Ceramic science and engineering as academic disciplines are concerned with the understanding of structure, properties, processing, and applications. Thus, ceramics is multidisciplinary learning from the principles provided by physics and chemistry, and the principles of chemical engineering, mechanical engineering, electrical engineering, etc. These principles are invoked through the continuum of ceramics processing, structure, and properties, with the chemical sciences playing a larger role in the processing and structure relationship and with physics and the other disciplines being more important in relating structure and properties.

There are 12 institutions certified by the Accreditation Board for Engineering and Technology that grant B.S. degrees in ceramics or ceramic engineering. There are 15 schools that offer formal advanced degrees in the field; many more schools granting degrees in "materials science" or "materials engineering" have faculty performing research on ceramics. Approximately 300 B.S., 90 M.S., and 35 Ph.D. degrees are awarded annually in ceramics or ceramic engineering.

At present, a large fraction of the B.S. graduates enter the more traditional areas of the ceramics industry such as glass and refractories. An increasingly larger fraction of the B.S. graduates perhaps 40 percent, currently enter the electronics field. On the other hand, virtually all the M.S. and Ph.D graduates are filling positions related to high-technology electronic and structural ceramics. The demand for advanced-degree graduates, particularly Ph.D.s, is far greater than the supply. As a result, many positions that could and should be filled by ceramics graduates are being filled by graduates from related fields, such as metallurgy, geology, and chemistry, who are being retrained during their first few years of industrial experience.

As high-technology ceramic production expands, the demand for B.S. graduates will increase. By 1990, most B.S. graduates will be entering the high-technology industries. At present, probably more than half of those entering graduate ceramics education have an undergraduate degree in the field. Thus, the number of B.S. graduates must grow to meet the needs of both expanding industry and expanding graduate research.

The livelihood of many of the undergraduate programs depends critically on the coexistence of a strong externally funded graduate program within the same department. A significant fraction of any new external support should be invested in those schools having both strong undergraduate and graduate programs to ensure that a continuous supply of high-quality B.S. graduates is available for both industry and graduate study.

Attracting new faculty and more students of high quality into ceramics programs in universities requires strengthening of these programs. Critical new efforts that permit the student to be a part of a balanced program supported by involvement of other disciplines and the use of modern equipment would be most important.

CONCLUSIONS AND RECOMMENDATIONS

The general conclusions reached in this study are:

- High-performance ceramics represent a class of materials that are essential for many electronic, optical, and structural applications in industry and defense. Such materials, which include monolithic ceramics, ceramic composites, and coatings, will find an expanded future use in such applications.
- The United States has led in basic research and development of products in high-

performance ceramics. However, this position is shifting because of the strong challenge to the United States from Japan and, to a lesser extent, from Europe. The primary aspect of this challenge is their ability to transfer fundamental research in high-performance ceramics to engineering practice and commercialization rather than their superiority over this country in basic research.

- Engineering research as well as basic research on high-performance ceramics must be strengthened and given needed impetus to ensure U.S. scientific, technological, and engineering leadership in the field. This calls for a large and appropriately coordinated effort on the part of universities, industry, and the federal and perhaps state governments.
- Preparation technology for improved thin films and layer structures designed for electronic and optical applications (e.g., substrates or capacitors), mechanical applications (e.g., friction and wear reduction), and chemical and thermal protection could greatly benefit from advances in ceramic processing and property science. The preparation science of layered ceramics and the study of surfaces, interfaces, and joints require interdisciplinary efforts by ceramists with other scientists and engineers.
- Completely new, multicomponent ceramic systems offer the potential of higher levels of performance as electronic or structural ceramics. Advances in chemistry, including chemical modeling and computation, offer improved predictive power as well as the ability to handle complex systems.
- The mechanical behavior of ceramics and tough composites for both electronic and structural use, including the behavior of ceramic fibers and interfaces, is an area of rapid progress. The identification of flaws as well as other high-temperature failure mechanisms needs further study, quantification, and the development of means to eliminate or control them. Toughness characterization, life prediction, testing, reliability assurance, and design with brittle materials must be improved.
- Interdisciplinary research offers major strategic gains on the complex problems of advanced ceramic processing and application, which must be solved as a materials system to supplement the knowledge gained by study of the isolated components. The components range from raw materials preparation, forming, sintering, joining, and fabrication to nondestructive evaluation, and include all aspects of design for current and future high-performance requirements for both electronic and structural ceramics applications. Novel chemical routes that combine or bypass stages of conventional ceramic processing also need in depth study.

No single university, government laboratory, or industrial organization has all the elements of what is foreseen as the necessary critical mass of multidisciplinary personnel to undertake the required programs. There are indeed several establishments that have such a grouping of talented personnel, and a very small number of others are in the early stages of development. However, neither any single U.S. organization nor the total of all pertinent U.S. organizations in these subjects currently has the dedication and focus of the Japanese efforts in these fields.

The following recommendations are offered for consideration:

- The United States should establish an initiative based on various programs designed to expand ceramic science and engineering research at universities. Such programs would be so configured as to ensure a reservoir of more trained graduate students, as well as a greater number of critical-mass size groups involved in ceramics research at universities. The bulk of the support for these programs could be provided by the National Science Foundation (NSF) and other federal agencies with the remainder (say 40 percent) coming from industrial, university, state, and other matching sources. A substantial portion of the funding in the early years should be for major equipment items. Several of these programs should be established to strengthen the ceramic education infrastructure at universities that already have existing strong undergraduate and graduate programs in ceramic science and engineering.

- A "summer institute" program should be established. In this program a number of leading ceramic researchers could gather for about a month and discuss their work, set priorities, make recommendations to NSF and other agencies for future funding, and coordinate efforts. Different groups concerned with different aspects of ceramics could rotate annually, with perhaps a core group of individuals. University, industry, and materials laboratory personnel would be included.
- The establishment of at least one large interdisciplinary ceramic center should also be considered. This center could be equipped to carry out research on the basic science and engineering of ceramic processing and the application of ceramics on a practical scale, including the relation of processing variables to specific model parts. For example, it could include experimental capabilities for detailed in-process studies of flows and reactions occurring during processing and with computer support (perhaps including access to a supercomputer) to permit detailed modeling of processing. Visits from scientists and engineers could be encouraged and provided for in the budget. The center could be viewed as a national facility for interdisciplinary experimental and theoretical work in electronic and structural advanced ceramics. It would contain an internationally recognized strong nucleus of permanent staff who would attract on a temporary basis individuals from industry, universities, and the national laboratories. Such a center would require funding, which could be jointly provided by federal agencies, industry, and perhaps state governments.

Report of the
Research Briefing Panel on Scientific Frontiers and the Superconducting Super Collider

Research Briefing Panel on Scientific Frontiers and the Superconducting Super Collider

Stanley Wojcicki (*Chairman*), Stanford University, Palo Alto, Calif.
Barry Barish, California Institute of Technology, Pasadena
Kenneth Fowler, Lawrence Livermore Laboratory, Livermore, Calif.
Barry H. Honig, Columbia University, New York, N.Y.
Arthur K. Kerman, Massachusetts Institute of Technology, Cambridge
C. Kumar N. Patel, AT&T Bell Laboratories, Murray Hill, N.J.
Chris Quigg, Fermi National Accelerator Laboratory, Batavia, Ill.
David Schramm, University of Chicago, Ill.
Roy F. Schwitters, Harvard University, Cambridge, Mass.
Sam B. Treiman, Princeton University, Princeton, N.J.
Steve Weinberg, University of Texas, Austin
Frank Wilczek, University of California, Santa Barbara
Harry Woolf, Institute for Advanced Study, Princeton, N.J.

Staff

Donald C. Shapero, *Staff Director*, Board on Physics and Astronomy
Robert L. Riemer, *Rapporteur*, Board on Physics and Astronomy
Helene Patterson, *Staff Assistant*, Board on Physics and Astronomy
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

Report of the Research Briefing Panel on Scientific Frontiers and the Superconducting Super Collider

INTRODUCTION

Elementary-particle physics, the science of the ultimate constituents of matter and their interactions, has undergone a remarkable development during the past two decades. A host of experimental results made accessible by the present generation of particle accelerators and the accompanying rapid convergence of theoretical ideas has brought to the subject an unprecedented coherence. This activity has raised fresh possibilities and set new goals for understanding nature. The progress in particle physics has been more dramatic and more thoroughgoing than could have been imagined only a dozen years ago. Many of the deep issues then current have been addressed, and many of the opportunities then foreseen have been realized. This progress has brought us to an intellectual turning point comparable to the synthesis of classical physics that preceded the discovery of relativity and quantum mechanics in the late nineteenth century. We are thus compelled by these special developments in science to build an unprecedented facility to provide the experimental basis for further progress in our understanding of the most basic forces of nature.

Experimental investigation of some of the fundamental questions in elementary-particle physics requires energies higher than those provided by any accelerators now in operation or under construction anywhere in the world. For this reason, the U.S. elementary-particle physics community is now preparing a proposal for a very high energy superconducting proton-proton collider, the Superconducting Super Collider (SSC). This major new accelerator complex would be based on the accelerator principles and technology that were developed in connection with the construction of the Fermilab Tevatron and on extensive work on superconducting magnets in the United States over the past 20 years. The proposed super collider would have an effective energy range about 60 times higher than any collider now in operation and 20 times the energy of the Tevatron collider now nearing completion. The SSC is needed to answer some of the pressing questions in elementary-particle physics. In addition, the new energy range that would become accessible with the super collider represents new and uncharted territory, the exploration of which can be expected to bring about a number of important consequences.

In what follows, we summarize our current understanding in this field. We describe recent successes, both theoretical and experimental, and we review the questions raised by the standard model. The super collider complex is subsequently briefly described. The main part of this report treats specific examples of experiments that could be performed with the super collider that may elucidate current questions in particle physics and cosmology.

HISTORICAL BACKGROUND

Elementary-particle physics is the study of the basic nature of matter, energy, space, and time. Elementary-particle physicists seek to understand the fundamental constituents of matter and the forces that govern their behavior. In common with all physicists, they search for the unifying principles and physical laws that determine the nature of the material world around us.

The goals of particle physicists are a natural continuation of a perpetual human quest to understand the world. The belief that the universe is rational, that matter is composed of relatively few simple constituents, and that the laws governing them are fundamentally simple and comprehensible to the human mind has pervaded the history of science in the western world. These ideas were confirmed by the theoretical and experimental advances achieved by scientists in the nineteenth century. The periodic table of Mendeleev, a systematic organization of the information obtained up to that time on the nature of matter, is in some sense a summary of their achievement. The modern phase of physics began at about the same time with the study of atomic emission and absorption lines in the visible spectrum and the discovery of the electron.

The atom, the atomic nucleus, and the elementary particles of which they are composed are too small to see or study directly. Throughout this century, physicists have devised ever more sophisticated detection devices to observe the traces of these particles and their constituents, and they have created increasingly energetic beams of particles to probe more deeply into the structure of matter. Early examples are the use of x rays to probe the electronic structure of the atom and the use of radioactive sources to study the atomic nucleus.

A fundamental law of quantum mechanics is that increasing the energy of probing particles allows us to probe matter on a finer scale. The progression to ever-smaller distance scales achieved in this century from atoms to nuclei to nucleons (protons and neutrons) and finally today to quarks and leptons depended on the attainment of ever-higher energies. Thus particle physics today is the natural descendant of the atomic physics and nuclear physics of yesterday.

To achieve the ability to probe smaller distances, it was necessary to go beyond the naturally occurring sources of energetic projectiles from radioactive sources. Utilizing the fact that protons and electrons (the constituents of ordinary matter) are electrically charged and experience forces in the presence of electric and magnetic fields, scientists have devised increasingly sophisticated methods of accelerating them to high-energies. The colliding of high-energy particles and the analysis of collision products is at the heart of experimental particle physics. For this reason, the field is often called high-energy physics.

There are analogies between the situation in particle physics today and that in chemistry at the end of the nineteenth century. Elementary constituents (quarks and leptons), playing the role once played by the elements, have been identified, and the many regularities observed among them have been systematized. There was no fundamental understanding in the nineteenth century of the structure of the periodic table; similarly, we lack today an understanding of the laws that determine the observed pattern among the quarks and leptons. Revolutionary new ideas (quantum mechanics) were required to elucidate the reasons for the regularity of Mendeleev's table. Only time and further experimental data will tell us if equally dramatic new insights will be needed to explain present particle-physics puzzles. Before addressing those questions, it would be useful to summarize our present understanding in more detail.

THE REVOLUTION IN PARTICLE PHYSICS

Forty years ago, ordinary matter was thought to consist of protons, neutrons, and electrons. Experiments probed the structure of these particles and explored the forces that bind them together into nuclei and atoms. In the course of these experiments, physicists discovered more than a hundred new particles, called hadrons, which had many similarities to protons and neutrons. None of these particles seemed more elementary than any other, and there was little understanding of the mechanisms by which they interacted.

Since that time, a radically new and simple picture has emerged as a result of many crucial experimental discoveries and theoretical insights. It is now clear that the proton, neutron, and other hadrons are not elementary. Rather, they are composite systems made of much smaller particles called quarks, much as an atom is a composite system made up of electrons and a nucleus. The existence of five kinds of quark has been established, and initial experimental evidence for a sixth species has been reported.

Unlike our view of the proton and neutron, our view of the electron as an elementary constituent of matter that is structureless and indivisible has survived the revolution intact. However, we now know that there are six kinds of electronlike particles called leptons. Both quarks and leptons appear to be grouped in three families of two members each. According to our present understanding, then, ordinary matter is composed of quarks and leptons.

The world around us is made of just three of those fundamental particles—the two lightest quarks and the electron. The other elementary constituents of matter are created under extreme conditions such as may exist in collisions of energetic particles either in cosmic rays, accelerator beams, or the interiors of exotic stars. They were also presumably present in abundance some 15 billion years ago in the early stages of the creation of the universe.

Nature contrives enormous complexity of structure and dynamics from the six quarks and six leptons now thought to be the fundamental constituents of matter. Physics seeks to simplify the description of nature by finding the underlying causes of natural occurrences and by relating apparently distinct phenomena. The result of this effort has been to show that all natural processes may be understood as manifestations of a small number of fundamental interactions. For half a century, physicists have recognized four basic forces: gravitation, electromagnetism, the weak interaction responsible for certain radioactive decays, and the strong force that binds atomic nuclei. An important difference between quarks and leptons is that one of these four interactions, the strong force that binds quarks together to form hadrons, does not affect leptons. Both quarks and leptons are acted on by the three other fundamental forces.

Over the past two decades, great progress has been made in understanding the nature of the strong, weak, and electromagnetic forces. The weak and electromagnetic forces have been unified by a theory whose predictions have been verified by many inventive experiments, the culmination of which was the discovery of the *W* and *Z* particles in 1983. These carriers of the weak force are analogs of the photon, the carrier of the electromagnetic interaction, whose existence was postulated early in this century and established experimentally by the middle 1920s. In addition, there is indirect but persuasive evidence for particles called gluons, the carriers of the strong force. The strong, weak, and electromagnetic interactions are all described by similar mathematical theories called gauge theories. At present, the role played by the gravitational force in elementary-particle physics is unclear. The effect of gravity on the behavior of elementary particles is so small that it usually can be ignored.

The experimental measurements and discoveries that shaped the revolution in particle physics were made possible by harnessing new accelerator and detector technologies that permitted the exploration of new energy domains in novel and incisive ways. Accelerator advances included the exploitation of the strong-focusing principle in synchrotrons; the creation of intense high-energy beams of neutrinos; the invention of colliding-beam accelerators (colliders) in which counterrotating beams of high-energy particles collide head on; the development of bright sources of nearly monoenergetic antiprotons; and the introduction of large-scale, energy-efficient, high-field superconducting accelerator magnets. Among the advances in observational techniques were the utilization of the bubble chamber for the observation of reaction products and the parallel development of a series of ever more capable electronic detectors; the mastery of fast digital electronics for data acquisition and processing; the evolution of methods for managing and analyzing vast quantities of data; and the construction of large, complex detector systems exploiting the capabilities of a variety of individual devices. Each sortie into a new energy regime, each improvement in our ability to search for rare processes, and each increase in sensitivity for their detection has led to new insights and, often, to the discovery of unexpected and revealing phenomena.

With the identification of quarks and leptons as elementary particles and the emergence of gauge theories as descriptions of the fundamental interactions, we possess today a coherent point of view and a single language appropriate for the description of all subnuclear phenomena. This development has made particle physics a much more unified subject, and it has also helped us to perceive common interests with other specialties. One important by-product of recent developments in elementary-particle physics has been a recognition of the close connection between this field and the study of the early evolution of the universe from its beginning in a tremendously energetic primordial explosion called the big bang. Particle physics provides important insights into the processes and conditions that prevailed in the early universe. Deductions from the current state of the universe can, in turn, give us information about particle processes at energies that are too high to be produced in the laboratory—energies that existed only in the first instants after the primordial explosion.

WHAT WE WANT TO KNOW

Developments in elementary-particle physics during the past decade have brought us to a new level of understanding of physical laws. This new level of understanding is often called the standard model of elementary-particle physics. As usual, the attainment of a new level of understanding refocuses attention on old problems that have refused to go away and raises new questions that could not have been asked before. The quark model of hadrons and the gauge theories of the strong, weak, and electromagnetic interactions organize our present knowledge and provide a setting for going beyond what we now know. A useful analogy can be made between our present-day understanding of elementary-particle physics and the situation in atomic physics in the early twentieth century or in the beta-decay studies in the 1930s.

The Bohr model of the hydrogen atom was a radical idea that initiated a revolution not only in the field of atomic physics but also in our whole philosophical approach to physics. By incorporating into one theory a new concept of quantization of angular momentum together with the classical ideas of electrodynamics, Bohr succeeded in creating a formalism that was remarkably successful in explaining the observed spectral lines of hydrogen. But the model had obvious difficulties. Not only did it fail to predict the spectra of more complicated atoms, but it also left unresolved the fundamental question of why the electron does not fall into the center of the atom by radiating away all its energy. Nevertheless, on the basis of its partial success, scientists believed that the ultimately correct theory would almost certainly contain some of the ideas embodied in the Bohr theory.

A similar situation existed in the case of Fermi's weak interaction theory. The scheme was invented to explain the phenomena in the million-electron-volt (MeV) energy domain that characterize the spontaneously occurring beta-decay processes. Its extrapolation to higher energies was remarkably fruitful, since it successively provided the appropriate framework to describe muon decay and capture, strange-particle decays, and, finally, neutrino interactions at energies of billions of electron volts (GeV). The accuracies of the experimental agreements were startling because it was apparent that the theory was only a low-energy approximation that broke down when taken to the domain of hundreds of GeV. Because of its successes, however, one was confident that elements of the Fermi theory would eventually become part of a more complete description.

Particle physics may be in a similar situation today. Although the standard model provides a framework for describing elementary particles and their interactions, its success prompts us to seek a more comprehensive understanding. For example, we do not know what determines such basic properties of quarks and leptons as their masses. Nor do we understand fully the origin of the differences between the massless electromagnetic force carrier (the photon) and the massive carriers of the weak force (the W and Z particles). Existing methods for dealing with these questions involve the introduction of many unexplained numerical constants into the theory—a situation that many physicists find arbitrary and thus unsatisfying. Physicists are actively seeking more complete and fundamental answers to these questions.

Another set of questions goes beyond the existing synthesis. For example, how many kinds of quark and lepton are there? How are the quarks and leptons related, if they are related? How can the strong force be unified with the already unified electromagnetic and weak forces?

Then there are questions related to our overview of elementary-particle physics. Are the quarks and leptons really elementary? Are there yet other types of forces and elementary particles? Can gravitation be treated quantum mechanically as are the other forces, and can it be unified with them? More generally, will quantum mechanics continue to apply as we probe smaller and smaller distances? Do we understand the basic nature of space and time?

Given this list of questions, it is not surprising that there are many directions of theoretical speculation departing from the current paradigm. Many of these speculations imply important phenomena at energies that are at present beyond our reach. Although theoretical speculation and synthesis are valuable and necessary, particle physics cannot advance without new observations. In the recent past, crucial observations have come from a variety of sources, including experiments at accelerators and nuclear reactors, nonaccelerator experiments (cosmic-ray studies and the search for proton decay), and deductions from astrophysical measurements. All our current ideas, embodied in the standard model, point to one trillion electron volts (1 TeV is an energy equivalent to approximately 1,000 proton masses) as the energy scale on which new phenomena can be expected. But a diversity of experimental initiatives will have to be mounted in order to explore thoroughly the new regime of energy and distance. A detailed examination of a great variety of conjectured extensions of the standard model shows that the super collider, with the wealth of different experiments that it would support, is the instrument of choice for exploring this new domain. At the same time, these extensions of the standard model set the parameters for the new accelerator.

The super collider will not foreclose all the other investigations in the field of particle physics. The accelerators currently under construction (like the electron-positron colliders at the Stanford Linear Accelerator Center [SLAC] and the European Organization for Nuclear Research [CERN], and the Tevatron proton-antiproton collider at Fermi National Accelerator Laboratory) will provide detailed quantitative tests of the standard model and its relation to cosmology. More sensitive proton decay experiments, better neutrino mass measurements, improved searches for monopoles in cosmic rays, and more precise searches for forbidden decay processes are examples of other important lines of non-SSC investigations that will continue to be pursued. However, according to our present knowledge of elementary-particle physics, our physical intuition, and our past experience, most clues and information will come from experiments at the highest-energy accelerators.

DESCRIPTION OF THE SUPER COLLIDER

The super collider would have two counterrotating beams of protons guided by superconducting magnets. Each beam would be accelerated to 20 TeV and the two beams would be allowed to collide. Six different collision points around the circumference of the main ring are proposed in current designs. Sophisticated detectors would be installed at the interaction regions, or collision points. Many crucial experimental tests have already been framed, and conceptual designs of several complementary detectors have been carried out.

We must note that it is not the collisions of protons on protons as such that are of primary interest. As discussed previously, protons are composite systems formed of quarks and gluons, the latter providing the binding force that holds the proton together. It is these "hard" collisions of a constituent of one proton with a constituent of the other that provide us with information about the fundamental interactions. Very roughly, each constituent carries about one tenth of the proton's total energy and thus colliding beams of 20-TeV protons are required to produce numerous constituent-constituent collisions at a few TeV. The ease with which the relatively rare hard scatterings of constituents have been observed in proton-antiproton collisions at the CERN collider (with a beam energy of about 0.3 TeV) gives us confidence that important physics results can be derived from the complex collisions that would occur in a super collider.

The proposed accelerator scheme makes use first of an injector system consisting of a linear accelerator followed by two circular accelerators. This system accelerates protons to about 1 TeV, at which point they are injected into the main ring for the final acceleration phase. The diameter of the main ring will be somewhere between 30 and 50 km, depending on the strength of the magnetic field in the magnets chosen for the super collider. Several different magnet styles are currently under consideration. The eventual choice of the magnet style will be made with the objective of minimizing the construction and operating costs of the machine and maximizing its reliability.

All the magnets under consideration use the superconducting technology and cryogenic systems first successfully employed on a large scale at the recently completed Tevatron ring at the Fermi National Accelerator Laboratory. Superconducting magnets make the super collider feasible by significantly reducing the required power consumption and enabling one to achieve higher field strengths than are offered by the conventional room-temperature magnets.

Current super-collider designs use proven technology on a vast scale. The linear dimensions of the super collider would be at least 15 times larger than those of the Tevatron and 3 to 5 times larger than those of the collider (LEP) currently under construction by CERN. While the Tevatron requires about 1,000 superconducting magnets, the super collider would employ perhaps as many as 14,000. This very large scale presents extremely challenging problems in such areas as manufacturing techniques, quality control, reliability, civil engineering, instrumentation and controls, communications, and installation and repair logistics. The magnitude of these problems is rather new to accelerator science and technology but manageable by an appropriate extension of present skills and experience. Creative solution of these problems calls for a new partnership between the basic research community and industry, which may bring important advances in technology of practical value. The detectors provide another great challenge to the experimentalists. Only a tiny fraction of the 10^8 interactions occurring every second will be of interest. The techniques necessary to identify and record these interesting events in a small fraction of a second will advance the frontiers of electronics and computer technology.

SCIENTIFIC QUESTIONS FOR THE SUPER COLLIDER

A major accelerator facility is not constructed to carry out a single experiment or measurement but rather to make possible a great diversity of investigations over the accelerator's lifetime, which is measured in decades. The evolution of the experimental program is guided by results of early experiments, by improvements in detector and accelerator technology, and by theoretical insights. Because of this, we cannot describe in advance the full scope of the research program for a super collider. We indicate, however, some of the issues to be addressed in the first round of experiments. These are representative of the questions that form the basis for the scientific justification for constructing a super collider and of the opportunities that a super collider would present.

The Origin of Mass

Imagine a universe pervaded everywhere by a uniform magnetic field. In such a universe, the motion of charged particles would appear to be rather complicated, because the particles' motions would be influenced by the universal magnetic field as well as by specifically applied forces. Eventually, as physics developed, some genius would realize that a simple law of motion (namely Newton's) is really the basic one and that all the complicated spiralings observed in his universe are caused by a pervasive background field.

The current standard model of the weak interactions suggests that a similar situation is realized in our universe. The field involved is not a magnetic field but rather what is called a Higgs field. The equations describing the weak interactions would take a simpler and more symmetrical form if there were no background Higgs field. One simplification that would occur is that the masses of the W and Z bosons and of many other particles would vanish. This simplification evidently goes too far for the real world. The best we can do is to postulate that the basic model operates in a world pervaded by a background Higgs field that partially hides its full symmetry and simplicity. The standard model, with this standard assumption, accurately describes a wide variety of observations—including the existence and mass of the W and Z bosons recently found at CERN—but requires the existence of one or more scalar particles associated with this background field.

Certainly one main area that we can hope to clarify with the super collider is the exact nature of the Higgs field and its interactions with other matter. Why is the super collider likely to be an appropriate tool? To see the answer, consider the reason that the Higgs field pervades our universe. It must be because the total energy density is minimized by having this field present. To change the magnitude of the field significantly, or make manifest one of its particles, we must supply enough energy density to overcome the natural tendency of the field to revert to its normal universal background value. We estimate that the energies that would be available at the super collider are almost certainly sufficient to do so.

Our present theoretical understanding of Higgs fields is primitive. There must be at least one such field, but there may well be many, each associated with its own particle. The super collider will open a window on this now dimly perceived sector of elementary-particle theory and help us to understand the reason for the apparently chaotic pattern of elementary-particle masses, caused, according to the present theory, by the interactions with the Higgs field or fields.

Families

The recent discovery of the third family of quarks and leptons has sharpened the problem of understanding the replication of particles in families. Three families of elementary particles are known that have, within a very well-defined sense, identical strong, electromagnetic, and weak interactions but different masses. Why does nature repeat itself in this way? Are there still more families? Do Higgs particles come in families? Are there other particles that do not fit into the same repetitious pattern? There are many such questions; they will only be answered by experiments at higher energies than those that we can now attain.

One theoretical approach to understanding the existence of families postulates the existence of symmetries under which the different families are interchanged. Given such symmetries, the existence of one family necessarily implies the existence of others. A particularly appealing version of this idea involves combining the family symmetry and the gauge symmetries of the strong, weak, and electromagnetic interactions into one gigantic, all-encompassing symmetry, hidden by the presence of suitable background Higgs fields. Implementing such symmetries could lead one to expect a new class of family interactions mediated by heavier analogs of the *W* and *Z* bosons that might be created at super-collider energies.

Chirality

There is a peculiar asymmetry in the weak interactions of the observed quarks and leptons. Simply put, the *W* bosons prefer to couple to these particles if they appear to be spinning clockwise, as viewed by an observer that they are approaching. We say that the weak interactions mediated by *W* bosons couple to left-handed quarks and leptons. It turns out that they couple to right-handed antiparticles (e.g., antiquarks or positrons). Many people have speculated that this asymmetry arises by a mechanism analogous to the one that causes the asymmetry between electromagnetic and weak interactions to appear in the standard model. According to this view, there will be either *W′* bosons that couple preferentially to right-handed rather than left-handed quarks and leptons, or new sets of quarks and leptons with right-handed couplings to the familiar *W* boson. In many models, the symmetry between left and right is broken by the same Higgs fields that spoil the fundamental weak-interaction symmetry. In such models, the masses of the *W′* and *Z′* bosons involved, or of the right-handed quarks and leptons, cannot be much larger than the known mass of the *W* and *Z*—which would make them observable at super-collider energies.

SUPERSYMMETRY

Recently there have been many theoretical investigations of the physical consequences of a radically new kind of symmetry called supersymmetry that implies the existence of integral-spin replicas or partners of particles of half-integral spin and vice versa. If it is relevant to physics at all, supersymmetry must be a broken symmetry—for example, there is definitely no particle having the mass and charge of the spin-1/2 electron that has integral spin. Such a particle would have been observed long ago if it existed.

There may be significant advantages to the view that supersymmetry is not too badly broken, so that many supersymmetry partners of known particles could be produced at super-collider energies. These advantages include resolution of the mathematical inconsistencies arising in the calculation of many physical quantities—specifically the masses of the Higgs particles.

The conjectured supersymmetry would lead to a Higgs boson mass less than 1 TeV/c^2 and yield supersymmetric particles with masses also less than about 1 TeV/c^2. Up to the present, there is no conclusive experimental evidence for these superpartners, even though there have been attempts to interpret some unusual events observed at CERN with large amounts of energy radiated in an invisible form in terms of supersymmetric models.

Significant progress has been made recently in building unified field theories that include gravity. These ambitious theories, which incorporate supersymmetry in a fundamental way, involve objects called superstrings; they will be confirmed, significantly constrained, or ruled out by experiments that elucidate the possible role of supersymmetry at super-collider energies.

DYNAMICAL SYMMETRY BREAKING

A second possible solution to the Higgs problem assumes that the Higgs boson is not an elementary particle at all but a composite object made of elementary constituents analogous to quarks and leptons. Although they would resemble the usual quarks and leptons, these new constituents would be subject to a new kind of strong interaction that would confine them within about 10^{-17} cm. Such new forces could yield new phenomena as rich and diverse as the conventional strong interactions, but on an energy scale a thousand times greater—around 1 TeV. The new phenomena would include a rich spectrum of new bound states akin to the spectrum of known hadrons. Again, there is no evidence yet for these new particles.

COMPOSITENESS

Violent collisions among quarks also provide a window on the possible internal structure of quarks. Some physicists feel that the known quarks and leptons are too numerous to be the ultimate elementary particles. If quarks are themselves composite, it should be possible to excite their internal structure in violent collisions. One sign of this kind of internal excitation would be spectacular multijet events quite unlike those anticipated in the standard model. The absence of such events and the agreement of the observed quark-quark scattering cross section with the standard-model predictions imply upper limits on the size of quarks. It will be possible with the super collider to look for quark substructure down to a distance of about 10^{-18} cm.

In summary, both general arguments and specific conjectures for resolutions of the Higgs problem imply 1 TeV as an energy scale on which new phenomena crucial to our understanding of the fundamental interactions must occur. The origin of electroweak symmetry breaking is only one of the important issues that define the frontier of elementary-particle physics. However, because of its immediate and fundamental significance it must guide our planning for future facilities.

Y 4.Sci 2:99/
100

READEX

US GOVERNMENT PUBLICATIONS (depository)
House of Representatives Comt on Science...
The future of science : hearing...99th Cong,
1st sess 1986

Dec 86-20148
7 of 7
p.0580-0595

Cosmology and the Super Collider

Over the past few years, cosmology and particle physics have become increasingly interwoven. To understand the physics that took place in the high-temperature, high-density early universe, one is forced to look at the physics of elementary particles. Similarly, the unified theories of elementary-particle physics have striking consequences at extremely high temperatures and energies. The only laboratory available to check these extrapolations of unified theories is the first instants after the big bang when extraordinarily high temperatures and densities were reached. In some ways, the two fields have become symbiotic. The super collider will be operating at energies far beyond those previously achievable in a laboratory and will simulate the conditions that prevailed when the temperature of the universe was about 10^{17} K. These conditions occurred about 10^{-15} second after the primordial explosion. Direct observations by optical telescopes are limited to events that occurred roughly 300,000 years after the big bang because the universe was opaque to photons at earlier times.

To reconstruct what occurred in the early universe, it is necessary to know the nature of fundamental interactions at high energies and the complete spectrum of elementary particles. In particular, the relics left over from those early times are of fundamental importance to cosmology. Any long-lived particle produced in the primordial explosion would survive and be an ingredient in the present-day universe.

One of the major issues in cosmology is to find the dark matter of the universe. Studies of the motion of stars within galaxies and of galaxies within clusters have established that these systems must contain a great deal of matter in addition to what is visible in the stars. This nonluminous matter may in fact account for the bulk of the mass in the universe. The properties that we impute to the dark matter depend on the character of the small density fluctuations in the early universe that grew into the galaxies and clusters observed today. According to current ideas on galaxy formation, the dark matter may be quite different from the ordinary (baryonic) matter of which we are made. Particle physics yields a mechanism for generating the primordial density fluctuations and provides candidates for the dark matter as well. Experimentation at the super collider will allow broad searches for new particles that may play the role of the dark matter.

In addition to the possibility of illuminating the question of dark matter of the universe, the super collider will clarify the structure and symmetry of the fundamental interactions and allow us to extrapolate with greater confidence back to early times. One of the most interesting recent developments in cosmology has been the suggestion that the homogeneity and isotropy of the universe were established in an early phase transition. According to this scenario, the universe began in a highly symmetric phase in which all the fundamental interactions were equivalent and evolved to the present low-symmetry phase in which different forces have different manifestations.

Another problem in astrophysics concerns unusual events produced in the infrequent collisions of very high energy cosmic rays with the Earth's atmosphere. Some of these events indicate unexpected phenomena occurring in the range of energies that can be explored systematically by the super collider. Such an exploration would make possible unambiguous interpretation of the hints provided by ultra high energy cosmic-ray events.

CONCLUSION

The advances of the past decade have brought us tantalizingly close to a profound new understanding of the fundamental constituents of nature and their interactions. The standard model based on quarks and leptons organizes our current knowledge and defines the frontier of particle physics at constituent energies of about 1 TeV and the frontier of cosmology at times of about 10^{-15} second. There we await new discoveries about the unification of the forces of nature, the patterns of the fundamental constituents of matter, and the origin of the universe.

Report of the Research Briefing Panel on Computer Vision and Pattern Recognition

Research Briefing Panel on Computer Vision and Pattern Recognition

Azriel Rosenfeld (*Chairman*), University of Maryland, College Park
J. K. Aggarwal, University of Texas, Austin
William L. Alford, Defense Mapping Agency/STT, Washington, D.C.
Norman Caplan, National Science Foundation, Washington, D.C.
Su-shing Chen, National Science Foundation, Washington, D.C.
Chao K. Chow, IBM Corporation, Yorktown Heights, N.Y.
Thomas M. Cover, Stanford University, Stanford, Calif.
Martin Fischler, SRI International, Menlo Park, Calif.
King-sun Fu, Purdue University, West Lafayette, Ind.
Robert Haralick, Machine Vision International, Ann Arbor, Mich.
Richard Heydorn, Johnson Space Flight Center, Houston, Tex.
Takeo Kanade, Carnegie-Mellon University, Pittsburgh, Pa.
Ernest Kent, National Bureau of Standards, Gaithersburg, Md.
Robert D. Leighty, U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Va.
George Nagy, University of Nebraska, Lincoln, Nebr.
Theodosios Pavlidis, AT&T Bell Laboratories, Murray Hill, N.J.

Staff

Gary E. Clark, *Staff Director*, Computer Science and Technology Board
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

Report of the Research Briefing Panel on Computer Vision and Pattern Recognition

1. INTRODUCTION

Computer vision is the process of deriving information about a scene by computer analysis of images of the scene. The images are obtained by a sensor, such as a television camera, and are converted to digital form for processing by a computer. The goal is usually the recognition of objects that appear in the scene. The analysis makes use, in part, of techniques from fields such as image processing, pattern recognition or classification, and artificial intelligence.

The field of computer vision is over 30 years old. As far back as the 1950s work was being done on such tasks as reading printed characters, counting blood cells, and recognizing military targets by analyzing digitized images. By the 1960s character reading had become a commercially successful application, and by the 1970s systems for medical applications such as blood counting had become commercially available. Research on military applications continued to be supported throughout this period, leading to successful demonstrations in areas such as tactical target detection and recognition and to practical use in terminal missile guidance. The late 1960s saw the beginnings of research on robot vision systems as well as on other industrial applications such as product inspection. By the late 1970s such applications began to reach the factory floor, and today there are more than 100 companies marketing industrial computer vision systems and devices.

As these examples indicate, computer vision has had many practical successes, but there are still limits on its capabilities. Some of these limits are due to inadequate computer power, but others are of a more fundamental nature. It has been estimated that adequate computer analysis of some types of complex, real-word scenes will require on the order of billions of computer operations per image. The computer power required to do this at television rates (30 images per second) is several orders of magnitude beyond what is available today at reasonable cost. The algorithms that will be needed for such complex analyses are not yet well understood—in part, because of the high computational cost of experimenting with such algorithms. More fundamentally, there is as yet insufficient theoretical basis for designing algorithms that can handle complex scenes; not enough is known about the mathematical description or modeling of classes of scenes.

The recent rapid advances in computer

hardware technology have made it possible to design special-purpose, massively parallel computer architectures suitable for the fast implementation of vision algorithms. The advent of VLSI (Very Large Scale Integration of digital circuits) and related hardware developments will greatly alleviate today's computer power limitations on vision systems. Efforts now need to be undertaken to alleviate the algorithmic and theoretical limitations. A significant research program is needed to stimulate advances in vision algorithms and in the theoretical foundations that underlie them. As successful algorithms are developed, it will be possible to design architectures that are matched to their computational requirements. Conversely, the availability of new architectures will make it possible to conduct significant experiments with complex vision algorithms and to learn more about their properties and about how to design them. The opportunities for major advances in computer vision over the next decade are exciting and challenging.

Section 2 of this report describes representative computer vision techniques designed for tasks of various degrees of complexity, and indicates the limitations on these techniques. Section 3 summarizes the state of the art in several major application areas, including remote sensing, cartography, reconnaissance, navigation, robotics, industrial inspection, and document understanding, and lists what appear to be the necessary conditions for successful applications of computer vision using today's technology. Section 4 outlines a set of research and development needs aimed at expanding the range of feasible applications. Finally, Section 5 discusses possible strategies for maximizing the payoff resulting from such a research program.

2. COMPUTER VISION TECHNIQUES

Before describing some typical computer vision techniques, we introduce some basic terminology. A computer vision system deals with images obtained from a *sensor* that views or scans the scene. In order to process these images by digital computer, they must first be converted into digital form, i.e., into arrays of numbers. Such an array is called a *digital image*, and its elements are called *pixels* (short for "picture elements"). The value of a pixel represents the brightness of the image in the neighborhood of a given point. This brightness, in turn, results from the radiation received by the sensor from the scene along a given direction. The number of pixels should be sufficient to adequately represent the resolution of the sensor. Conventional television images are usually digitized into arrays of about 500 by 500 pixels.

2.1 Pixel Classification

In remote sensing of the earth's surface or the atmosphere from satellite altitudes, a single pixel represents a significant area, usually tens of meters in diameter, on the earth's surface. Typically, the sensor produces a number of images using different spectral bands in the visible and near infrared. When these images are digitized, we thus have a set of values associated with each pixel.

Such multispectral images have been extensively used for terrain or crop classification. The standard approach is to assign each pixel to a class on the basis of its set of spectral measurements, using statistical decision techniques.

The accuracy with which this type of classification can be carried out depends on how accurately we model the problem domain. If we simply characterize the classes by their spectral measurements, we find that these measurements are highly variable. As a result, classification based on statistical decision methods will have a high inherent error rate. The error rate is further increased by the fact that the ground area corresponding to a pixel may contain a mixture of classes. Improved results can be obtained by using models that incorporate physical information about the scene.

2.2 Two-Dimensional Object Recognition

In most applications of computer vision, the objects to be recognized are much bigger than a single pixel. We consider first the situation where the scene is essentially two-dimensional, for example, where the objects are characters on a document, lines on a drawing, or connectors on a circuit board. In this situation, any objects present in the scene should be fully visible on the image.

If the objects to be recognized are exactly known—for example, if they are characters from a known typeface—a straightforward "template matching" process can be used to recognize them. The main limitations on recognition accuracy are the precision with which the object shapes are known and the fidelity with which the image represents the scene; accuracy decreases if the image is degraded (has low contrast, is noisy or blurred) or if the objects themselves are imperfect (e.g., badly printed characters). We can optimize system performance if we have good models for both the object and the degradation processes.

In most situations, object shapes are not known exactly, and recognition accuracy is limited not only by image degradations but also by how well we can define (i.e., model) the shapes belonging to a given object class. A classical approach is to characterize the objects by a set of geometrical measurements or features (e.g., area, perimeter, number of holes) and to use statistical decision methods. If the classes to be recognized are sufficiently dissimilar, this approach can yield acceptable levels of recognition. Of course, its success depends on our ability to extract ("segment") the objects from their background so that the features can be measured correctly. Successful segmentation depends in turn on being able to characterize (i.e., model) how the objects differ from the background (e.g., characters are usually darker than the surface on which they appear).

For complex objects, the feature-based approach is usually not adequate to capture the subtle differences that distinguish the classes. Rather, it becomes necessary to model the objects in greater detail, e.g., as composed of parts that have given features and that are arranged in given configurations. (For example, we can model characters as composed of strokes that have specified shapes and are combined in particular ways.) To apply this "structural" approach, it is necessary to extract the parts from the image, measure their features, and determine what relationships hold among them. This information can be represented by a graph structure, and recognition can be achieved by matching this structure with stored graph structures representing the models for the given classes of objects. (Note that graph matching may involve searching through a very large space of possible matches.) This approach is much more flexible than feature-based techniques, but it may require models that are complex and difficult to specify adequately.

2.3 Three-Dimensional Object Recognition

Computer vision becomes much more difficult when it is necessary to deal with truly three-dimensional scenes—for example, when a robot has to recognize mechanical parts in a pile, or plan a path across a cluttered factory floor. Two factors make such three-dimensional tasks more difficult:

a. The image shows only a two-dimensional projection of the scene; the visible surface points in the scene lie at different distances from the sensor, but the position of a point in the image tells us only along what direction in the scene the corresponding surface point lies, not how far away it is.

b. Even if we knew the distances, we would only have information about the surfaces that are visible from the sensor; we cannot see the hidden sides of objects, or the parts of objects that are hidden by other objects.

A class of techniques has been developed over the past 15 years to deal with the first problem; these techniques are known collectively as "shape from . . ." methods. They infer information about the three-dimensional shapes of the visible surfaces in the scene from a variety of clues present in the image: brightness variations ("shading"), changes in the sizes and spacings of local patterns ("texture"), etc., by making use of assumptions about (i.e., models for) the shapes and surface properties of objects and the sources of illumination. The ultimate goal of these techniques is to determine the illumination, reflectivity, and spatial orientation of each visible surface point; these are the factors that give rise to the brightness of the corresponding image point. In effect, we are attempting to solve the "inverse optics" problem; if we know the values of these factors over the scene, it is easy to compute the image brightness, but the reverse is much harder. If we can estimate the values of these factors at each pixel, we now have a set of values, rather than a single value (brightness) associated with the pixel. In other words, we have replaced the original image by a set of images. These are known as "intrinsic" images, since each of them represents an intrinsic physical property computed at each visible point of the scene. The process of estimating the intrinsic image information is known as "recovery" (short for "recovery of intrinsic scene characteristics from an image").

Extraction of objects from an image of a three-dimensional scene becomes much easier if intrinsic image information is available. An object and its background may not be easily distinguishable on the basis of their brightnesses in the image; even if they have different reflectivities, variations in illumination and surface orientation may cause their brightness ranges to overlap. However, if we know the illumination and orientation, we can correct for their effects.

Even if we have intrinsic image information, recognition of three-dimensional objects is still difficult, since an object may be partially hidden by other objects. Even if it is not hidden, it may occur in the scene in many different orientations, so that its projection in the image may have many different shapes. One approach to recognizing objects in arbitrary orientations is to use a large set of two-dimensional shape models, one for "every" orientation. A more flexible approach makes use of a set of three-dimensional models for object parts and is based on an "inverse projective geometry" technique. Given a region in the image, not all of the possible three-dimensional object parts can project into it, and for those that can, not all orientations are possible. Suppose we find a set of regions all of which could be the projections of parts of the same three-dimensional object, all oriented consistently; then we have strong evidence for the recognition of that object.

2.4 OTHER TECHNIQUES

In this section we briefly mention some aspects of computer vision that were omitted from the previous sections for the sake of simplicity.

Local feature detection A region in an image may have a wide range of brightnesses, but it may still be distinguishable from its background because there is a sharp brightness discontinuity along its border. Many techniques for detecting such brightness "edges" in images have been developed, based on models for various types of physical discontinuities in the scene. There are also techniques for detecting special types of local image features involving brightness discontinuities—for example, thin lines or curves that contrast locally with their backgrounds.

Texture analysis A region in an image is sometimes distinguishable from its background because it has a characteristic pattern of brightnesses or local features. Many techniques have been developed for modeling and discriminating among such patterns, or "visual textures," and using them as a basis for segmenting an image. As mentioned earlier, texture variations across a region in an image can also be a useful clue to the three-dimensional orientation of the corresponding surface in the scene.

Stereo If two images of a scene taken from different (known) positions are available, and we can identify corresponding points (arising from the same scene point) in the two images, we can determine the three-dimensional location of the scene point by triangulation. The images are called a stereopair, and the triangulation process is called stereoplotting. The difficult problems are (1) reliably identifying corresponding pairs of points and (2) interpolating plausible three-dimensional surfaces based on a sparse set of point correspondences.

Motion Given a sequence of images of a scene taken at closely spaced times from a moving sensor, by comparing successive images we can derive information about the three-dimensional structure of the scene. Many techniques have been developed for measuring the "optical flow" of intensities in a time sequence of images and for inferring scene structure from the flow field.

Controlled illumination Control of illumination can make computer vision tasks easier in a variety of ways. It can be used to improve the contrast between an object and its background, e.g., by silhouetting the object. Three-dimensional information about a surface can be unambiguously derived by comparing the intensity variations across the surface when it is illuminated from several known directions; this method is called "photometric stereo." Patterned illumination can also provide three-dimensional surface information with the aid of triangulation techniques; this is known as "structured light" range measurement.

3. APPLICATION AREAS

3.1 Remote Sensing

Pixel classification based on spectral measurements, as described in Section 2.1, has been extensively used for crop classification and crop acreage estimation. The discussion here deals primarily with measurements obtained by satellite-borne multispectral scanners, but there are other applications that make use of color aerial photography. An important extension is to use several multispectral images taken at different times during the growing season.

Characterizing crop classes by their spectral measurements is not entirely adequate, since these measurements are sensitive to variations in soil background and planting data, as well as to atmospheric effects. Typical accuracies obtained, e.g., for discriminating corn from soybeans or winter grains from other crops, are only about 70 percent. Better results, with accuracies of about 85 percent, are obtained by introducing models for the growth cycle of the crops. We now represent the data by a set of estimated model parameters at each pixel; this transformation of the data often makes the classification problem easier by emphasizing parameters that best discriminate crops and ignoring other factors.

Further improvement should be achievable by introducing more detailed models that take into account the physical structure of the ground cover. One can model the scattering of radiation from vegetation in terms of parameters such as leaf angle distribution, leaf transmittance, and leaf reflectance, as well as the geometric placement of the plants. Of course, such models are quite complex, and extensive studies will be needed to define them quantitatively for a given class of crop scenes. The observed measurements often result from a mixture of several models, and it will also be necessary to develop general methods of handling such mixtures.

3.2 Cartography and Reconnaissance

Map makers use stereopairs of aerial photographs as a primary source of information about the earth's surface, including both measurement of terrain height and recognition of surface features. Systems proposed for operational use may incorporate some degree of automated stereoplotting, but are still highly dependent on interaction with human operators. By the 1990s, operational systems may incorporate computer vision capabilities for automated extraction and recognition of various feature types, such as hydrological features, vegetation, roads, and certain types of structures. In relatively flat terrain viewed from high above, feature recognition is essentially a two-dimensional problem, and can usually be handled by the methods described in Section 2.2.

Airborne reconnaissance is conducted in two modes: searching and monitoring. In the searching mode, targets can occur anywhere and the task is to detect and recognize them. The monitoring mode involves periodic examination of a known area to check on its operational status. In both modes, today's systems can provide computer assistance to a human operator. Selected types of targets, e.g., vehicles in infrared images, can be searched for, detected, and their positions signaled to the operator (which is known as "target cueing"). Systems now in the research and development stage are expected to achieve reliable target recognition capabilities through the use of contextual information.

3.3. Navigation

Terminal homing is the process of guiding a missile to a target by correcting its path with the aid of computer vision techniques. This generally involves matching a stored template of the target with a sensor image. Such techniques are used in the Pershing missile system. A proposed improvement will make use of a set of templates corresponding to different target ranges.

Mid-course guidance is the process of determining the location of a missile in order to correct its course. Present cruise missiles do this by matching a stored pattern of terrain elevations with radar altimeter data. A more robust approach would match stored templates with a high-resolution sensor image; this should be feasible now or in the near future.

There has been recent interest in the development of an autonomous land vehicle capable of traversing a road network, and ultimately of navigating across various types of terrain. Such a vehicle will make extensive use of computer vision capabilities for road detection and following, obstacle detection and avoidance, and detection of landmarks for navigation purposes. The Defense Advanced Research Projects Agency (DARPA) has initiated work on such a system under its Strategic Computing Program. Autonomous land vehicles, or mobile robots, will have many nonmilitary uses in areas such as construction, shipbuilding, and agriculture.

3.4 Industrial Inspection and Robot Vision

Most of today's industrial computer vision systems are designed to perform inspection tasks. Such tasks can be reliably automated, using today's technology, if the following conditions hold:

a. The object to be inspected has a regular or known shape, having sharply defined, high-contrast features;
b. Objects do not overlap, and have known orientations;
c. The lighting is controlled and uniform;
d. Defects are large and clear cut.

An example of an application area for which many working systems have been developed is printed circuit inspection.

Another important industrial application of computer vision is in robot vision systems. Most systems in current commercial use employ very simple techniques in highly constrained situations. A typical application is seam tracking in robot welding, using structured light triangulation techniques. Other systems operate on high-contrast images by first segmenting them into objects and background; this enables the system to deal with one-bit pixel arrays in which 1's are object pixels and 0's are background pixels. If the segmentation can be done reliably, perhaps with the aid of controlled lighting, such "binary" vision systems can operate at high speeds. One application for such systems is the lifting of objects from a conveyor belt, where the objects can have only a few stable positions. Picking objects out of a bin is a much more difficult task, since the objects can be in arbitrary three-dimensional orientations and the lighting inside a bin cannot be completely controlled.

Robot vision will find its greatest use in situations where the timing, placement, orientation, and identity of arriving objects are not completely constrained. Ultimately, robots will need to recognize and locate specific objects in piles of mixed objects. Techniques have been developed that can handle such tasks, but they are still computationally costly and slow.

3.5. Document Understanding

Document understanding started with typewritten optical character recognition and has been expanded to include recognition of typeset copy, engineering drawings, maps, and other graphical information. While there is no "universal" machine for reading arbitrary documents, there have been many partial successes.

The main factor influencing success in this area has been control of the input. Chronologically, the problems solved were specially shaped numerals on checks; bar codes; monospaced, single-font readers; readers for a limited number of fonts; and reading of carefully formed hand-printed letters and numerals within boxes.

A second factor has been the use of prior knowledge or context. For example, in postal sorting of outgoing mail the postal code is compared with the city and state and a decision is made only if both agree. Similarly, in scanning line drawings of logic circuits, the graphic input is checked against circuit constraints.

Areas ripe for near-term development include multifont typeset text, typeset text interspersed with graphics, and graphics labeled with typeset text. An important point in general document processing is that it is not sufficient to classify the symbols. It is also essential to represent the geometric relations between them, since these too convey meaning.

3.6. Characteristics of Successful Applications

In each of the domains discussed above, many tasks can be successfully carried out using today's computer vision technology, but many other tasks still present research challenges. The following are some of the criteria that characterize today's successful applications:

a. Computer vision is easier in a restricted scene domain, in which the class of scenes can be modeled with sufficient accuracy. The domain can sometimes be simplified by controlling the scene environment, for example, by controlling the lighting.

b. Two-dimensional scenes are easier to deal with than three-dimensional scenes.

c. There is a trade-off between the time available to carry out a task and the computer power required. If a task must be performed under very tight time constraints, its implementation may require impractical amounts of computation.

An example of a task that is extremely difficult in terms of all these criteria is autonomous vehicle navigation. Here the scene domain is potentially unrestricted; the environment is uncontrolled; the scene is strongly three-dimensional; and the vehicle must navigate in real time.

In order to extend computer vision capabilities to domains that do not satisfy the above criteria, a substantial research and development effort will be necessary, as discussed in the next section.

4. RESEARCH AND DEVELOPMENT NEEDS

4.1. Sensors

Most of the discussion in previous sections has assumed that the sensor is some type of television camera, but many other types of sensors have been used in computer vision systems. The airborne and satellite-borne scanners used in remote sensing detect radiation in a set of spectral bands in the visible and near infrared. Thermal infrared imaging sensors that provide temperature information are often used in target detection applications. Sensors that operate in other spectral bands can also be used; for example, extensive research has been done on the application of computer vision techniques to x-ray images, both for medical purposes (for example, analysis of heart shape or detection of tumors) and for industrial inspection. One can also work with sensors that make use of nonelectromagnetic energy; for example, ultrasonic arrays are used for obstacle detection in various applications.

An important class of sensors provides three-dimensional information about a scene, rather than just intensity information. Radar sensors directly measure the range to every visible scene point, as well as providing information about the dielectric properties of surfaces. Conventional radars use microwave radiation, but systems using visible light have also been developed. The use of such sensors greatly reduces the difficulty of handling three-dimensional scenes, but it still does not provide information about hidden parts of objects. Under the proper circumstances, complete three-dimensional information about the scene can be obtained using tomographic reconstruction techniques, which yield a three-dimensional array of "pixels" representing "density" information at each point of a three-dimensional volume. Some work has been done on extending computer vision techniques to handle such three-dimensional arrays of data.

As we have seen, successful design of a computer vision system depends on adequately modeling the scene, the objects present in it, and the process by which the images are obtained. Most existing computer vision techniques were designed for images obtained by optical or near-infrared sensors. For sensors that make use of other types of radiation, such as x rays or ultrasound, the imaging process is quite different, and the computer vision techniques used should be based on different models. These remarks apply even more strongly to sensors that provide direct range information or complete three-dimensional information. As such sensors begin to be used in computer vision systems, new techniques will have to be developed based on the appropriate models.

In many situations, more than one sensor will be needed in order to carry out a given task, since different sensors will provide different types of information. This leads to the need to develop techniques for integration or "fusion" of information from different sensors. Registering the data obtained by different types of sensors is a nontrivial task, since the sensors generally have different resolutions and often have different imaging geometries.

Even when conventional sensors are used, the choice of sensor parameters may significantly affect the difficulty of computer vision tasks. There is an important engineering trade-off between sensor cost and computational cost. In some applications, for example in the document processing area, it has been shown that a moderate increase in sensor resolution and linearity made it possible to avoid a massive increase in computation.

4.2. Computers

Digital images contain voluminous amounts of data—typically, hundreds of thousands of pixels. Typical computer vision tasks may require hundreds of computer operations per pixel, for a total of tens or hundreds of millions of operations per image. If images need to be processed at television rates (30 per second), conventional computers are not fast enough.

The solution to this problem lies in the use of massive multiprocessing. The match between the computational needs of computer vision and the emerging computational capabilities of multiprocessing systems appears to be excellent. Computer vision tasks often involve the performance of identical computations, or sequences of computations, at every point in an image. Multiprocessing makes it possible to carry out such tasks in parallel by assigning different regions of the image to different processors. These processors can operate relatively independently; for many operations, only local communication between neighboring processors is needed. Long sequences of computations can be carried out by "pipelined" multiprocessors, with each processor operating on the output produced by the preceding one.

Major research and development efforts are under way worldwide to build multiprocessor systems for computer vision and image processing. Dozens of such systems have been designed and built in the United States, Japan, and Western Europe. Experiments with these systems will provide us with useful information about the types of multiprocessor architectures that will be most useful for computer vision.

The recent advances in VLSI make it possible to produce powerful special-purpose processors in quantity. A wide variety of different processor architectures can be used to speed up computer vision processes. It is now possible to design and fabricate chips that implement specific image processing, pattern recognition, and computer vision techniques, for example, template matching. By combining such chips very fast vision computers could be produced.

Since computer vision systems usually employ optical sensors, some types of initial processing of the sensor data can be done at very high speeds, in parallel, using optical filtering techniques. Optical and hybrid optical-digital methods can be used to perform a variety of useful operations on images.

4.3 Models

We have emphasized earlier that to achieve high levels of performance, computer vision systems should make use of good models for the imaging process and for the classes of scenes to be analyzed. Many types of scene models are highly specialized (e.g., models for crop/leaf canopies), while others are of a more general nature (e.g., Markov models for textures, fractal models for surfaces, reflectivity models). In some areas, good modeling techniques do not as yet exist; better methods are needed for modeling classes of complex object shapes (in two or three dimensions) or spatial relations between objects. More intensive work is needed on developing and improving models for commonly analyzed classes of scenes.

Models can provide a theoretical basis for designing computer vision systems. They make it possible to design optimal techniques for such problems as image segmentation, local feature detection, texture discrimination or recovery of intrinsic scene information, optimal sets of features for two- or three-dimensional shape recognition, and so on, and to predict the performance of these techniques. It should be pointed out that many of these problems are underconstrained or "ill-posed," and can be solved only if some type of model for the scene domain is specified.

In nearly all existing computer vision systems, models are used implicitly as a basis for designing the analysis techniques used by the system. An important current research area is the design of "expert" computer vision systems in which the models (i.e., the system's knowledge about the scene, the sensor, and other factors) is represented explicitly. By making this information explicit, it becomes possible to design the system to evaluate its own performance, determine the possible causes of errors, and take appropriate action. The addition of such capabilities to computer vision systems would lead to substantial improvements in their performance.

4.4. Computational and Algorithmic Tools

Computer vision systems make use of many computational techniques and tools that are also used in other areas of pattern recognition, artificial intelligence, and computer science. Improvements in these tools would be of great benefit to computer vision research. The following are some examples of techniques that are of particular potential importance:

a. Pattern matching techniques, including matching of images and of graph structures derived from images;

b. Methods of combining evidence obtained from multiple sources and of resolving conflicts;

c. Methods of effectively using context in decision making and recognition;

d. Techniques for controlling potential "combinatorial explosions" in the size of search spaces;

e. "Indexing" techniques for rapidly selecting, from a large data base, the models that are relevant to the current task;

f. Methods of partitioning problems into coherent components that can be analyzed independently;

g. Methods of communicating information between processes that are operating in parallel;

h. Knowledge acquisition techniques for building and organizing models or knowledge bases for given problems;

i. Techniques for effective human-machine communication to discuss iconic and spatial data.

4.5. Theoretical Foundations

Many of the specific processes used in vision systems have well-developed theories; for example, the techniques for recovering intrinsic scene properties from an image have a well-understood mathematical basis. In many parts of the field, however, ad hoc techniques are still extensively used. The development of theoretical foundations in other parts of the field would provide a basis for designing optimal algorithms, conducting controlled experiments, determining their statistical validity, and evaluating their results.

A more challenging task would be that of developing a general theory of the overall computer vision (or pattern recognition) process. A possible basis for such a theory lies in the concept of complexity. Kolmogorov has defined the complexity of an object as the length of the shortest computer program for describing that object. This notion of complexity turns out to be independent of the particular computational model that is used; it is a universal notion. The goal of image description is to derive just such concise descriptions for a given image. More precisely, given a language in which such descriptions can be formulated—i.e., a language for describing the given class of scenes—the goal is to find "good" or "sufficient" descriptions, in that language, for a particular image. This formulation is conceptually important even though, in real-world cases, it may be computationally impractical to realize. (The computation time required to find the shortest description of an object is, in general, unbounded.) Nevertheless, it would be desirable to assess the current ad hoc techniques in relation to the domain of theoretically achievable techniques—in particular, to assess how far away they are from the theoretical optimum.

In general, the field of computer vision can benefit from research in other related areas, including artificial intelligence, automatic inference, computational complexity, nonparametric modeling, and statistical pattern recognition. Computer vision researchers can also benefit from reviewing what is known about biological visual systems, since these systems constitute an existence proof that complex visual tasks can be carried out rapidly and reliably. The brain's ability to recognize objects in complex scenes in a fraction of a second demonstrates the power of massively parallel computation. Conversely, the design of special computer architectures for real-time computer vision may provide clues about the structure of the brain, and computer vision algorithms may suggest models for biological visual processes.

5. RESEARCH STRATEGIES

In this section we discuss some alternative strategies for supporting the research that will be needed to achieve substantial advances in computer vision. Such support will have to come primarily from federal sources, since it represents a long-term investment. Computer vision activities in the private sector have been confined almost exclusively to short-term product developments having predictable payoffs and involving relatively simple applications, in areas such as character recognition and two-dimensional industrial inspection.

The conventional approach to the support of basic research in a given area has been to fund a number of relatively small individual grants addressing specific aspects of the subject. This is the approach followed by the National Science Foundation, for example. As applied to computer vision, it has produced, and will continue to produce, important advances in the underlying theory, paradigms, principles, and techniques.

DARPA has adopted a different strategy in connection with its Strategic Computing Program. It has defined a number of major application goals, the achievement of which will require substantial advances on many fronts. In the area of computer vision, the application is embodied in the Autonomous Land Vehicle project, which will accomplish a series of milestone demonstrations involving road following, obstacle avoidance, and route planning over the coming years. The general strategy of this approach is to pick an application that is beyond today's state of the art, but that requires well-defined, specifiable advances. Other computer vision applications could be defined that would serve similar purposes; examples are extraction of cartographic features from aerial imagery, robot vision for outdoor construction tasks, and reading of unconstrained cursive handwriting.

Computer vision systems are made up of many components; they make use of image processing algorithms, data structures, and knowledge representations of many different types. Creating a computer vision system to perform a complex task such as vehicle navigation or extraction of cartographic features requires an extensive programming effort involving a team of researchers over a period of months or even years. It also requires major efforts in system integration, an area in which there is little past experience. The level of effort increases still further if the systems must be implemented on a new computer architecture for which a software environment for program development does not yet exist. Building complete vision systems of such advanced types thus can be done only by large research groups and requires substantial funding.

System building efforts can also be undertaken by cooperation among smaller research groups, but there are practical difficulties with this approach. It is often difficult to get researchers to agree on standards for software compatibility, for communication between processes, or even for data formats. Some attention should be given to developing mechanisms that would facilitate and encourage joint cooperative efforts. Competitions between teams of researchers taking different approaches to the same problems would also provide an incentive to produce high-quality results. The competition concept can be applied not only at the level of complete vision systems, but also in connection with smaller research efforts involving specific components of a vision system.

The most effective way to create large experimental vision systems is at a large, well-established research center. Most of the existing centers are associated with universities, but there are several government and corporate laboratories that are also well qualified in this area. It would be desirable to create additional centers of excellence in computer vision, so that experimental system building efforts in a number of different application areas could be pursued. Such centers could also serve as resources for smaller groups. This is particularly important for access to experimental computer systems; ordinarily only a few copies of such systems will exist, and it would be desirable to allow many groups to have access to them.

Computer vision researchers are much sought after by recruiters who want to use their expertise in state-of-the-art applications. It is important to provide incentives that will allow many of these researchers to remain in an academic environment, where they can be free to pursue long-term research goals and where they can also help in the training of the next generation of researchers.

The United States is still at the forefront of computer vision research, but other countries are not far behind. Japan has funded a series of major projects that have had significant impact in the areas of computer vision and pattern recognition. It is important that the United States take similar steps to maintain its lead in the vision field, which has potential payoffs in so many different areas of advanced automation.

○